Estimation of Growth and Extinction Parameters for Endangered Species

Brian Dennis; Patricia L. Munholland; J. Michael Scott

ESTIMATION OF GROWTH AND EXTINCTION PARAMETERS FOR ENDANGERED SPECIES[1]

BRIAN DENNIS

*Department of Forest Resources and Department of Mathematics and Statistics,
University of Idaho, Moscow, Idaho 83843 USA*

PATRICIA L. MUNHOLLAND

Department of Mathematical Sciences, Montana State University, Bozeman, Montana 59717-0001 USA

J. MICHAEL SCOTT

*United States Fish and Wildlife Service, Idaho Cooperative Fish and Wildlife Research Unit,
College of Forestry, Wildlife, and Range Sciences, University of Idaho,
Moscow, Idaho 83843 USA*

Abstract. Survival or extinction of an endangered species is inherently stochastic. We develop statistical methods for estimating quantities related to growth rates and extinction probabilities from time series data on the abundance of a single population. The statistical methods are based on a stochastic model of exponential growth arising from the biological theory of age- or stage-structured populations. The model incorporates the so-called environmental type of stochastic fluctuations and yields a lognormal probability distribution of population abundance. Calculation of maximum likelihood estimates of the two unknown parameters in this model reduces to performing a simple linear regression. We describe techniques for rigorously testing and evaluating whether the model fits a given data set. Various growth- and extinction-related quantities are functions of the two parameters, including the continuous rate of increase, the finite rate of increase, the geometric finite rate of increase, the probability of reaching a lower threshold population size, the mean, median, and most likely time of attaining the threshold, and the projected population size. Maximum likelihood estimates and minimum variance unbiased estimates of these quantities are described in detail.

We provide example analyses of data on the Whooping Crane (*Grus americana*), grizzly bear (*Ursus arctos horribilis*) in Yellowstone, Kirtland's Warbler (*Dendroica kirtlandii*), California Condor (*Gymnogyps californianus*), Puerto Rican Parrot (*Amazona vittata*), Palila (*Loxioides balleui*), and Laysan Finch (*Telespyza cantans*). The model results indicate a favorable outlook for the Whooping Crane, but long-term unfavorable prospects for the Yellowstone grizzly bear population and for Kirtland's Warbler. Results for the California Condor, in a retrospective analysis, indicate a virtual emergency existed in 1980. The analyses suggest that the Puerto Rican Parrot faces little risk of extinction from ordinary environmental fluctuations, provided intensive management efforts continue. However, the model does not account for the possibility of freak catastrophic events (hurricanes, fires, etc.), which are likely the most severe source of risk to the Puerto Rican Parrot, as shown by the recent decimation of this population by Hurricane Hugo. Model parameter estimates for the Palila and the Laysan Finch have wide uncertainty due to the extreme fluctuations in the population sizes of these species. In general, the model fits the example data sets well. We conclude that the model, and the associated statistical methods, can be useful for investigating various scientific and management questions concerning species preservation.

Key words: California Condor; conservation biology; diffusion process; endangered species; exponential growth; extinction; grizzly bear; inverse Gaussian distribution; Kirtland's Warbler; Laysan Finch; lognormal distribution; Palila; parameter estimation; Puerto Rican Parrot; stochastic differential equation; stochastic population model; Whooping Crane; Wiener process.

INTRODUCTION

The extinction of a population is a chance event. A population's growth inevitably displays stochastic fluctuations due to numerous unpredictable causes. Consequently, a species with an average negative growth rate might temporarily prosper, while a species with a positive rate might become endangered. The field of conservation biology has recognized the importance of accounting for stochastic factors in species preservation efforts (Shaffer 1981, Samson et al. 1985, Soulé 1986, 1987, Burgman et al. 1988, Lande 1988, Simberloff 1988). Such accounting in practice has proved no easy task.

Many mathematical models of population growth incorporating stochastic fluctuations have been studied; Nisbet and Gurney (1982) and Goel and Richter-Dyn (1974) provide excellent surveys. The chance of extinction or waiting time to extinction is frequently the focus of theoretical or numerical analyses of stochastic growth models (Capocelli and Ricciardi 1974, Richter-Dyn and Goel 1974, Feldman and Roughgarden 1975, Keiding 1975, Leigh 1981, Tier and Hanson 1981, Ginzburg et al. 1982, Braumann 1983a, b, Wright and Hubble 1983, Strebel 1985, Goodman 1987, Lande 1987, Iwasa and Mochizuki 1988, Lande and Orzack 1988, Dennis 1989a). Ginzburg et al. (1982), in particular, described the concepts of risk analysis as a framework for such extinction studies. These studies have tended to use relatively simple stochastic models such as univariate birth–death processes or diffusion processes, since the analyses are considerably easier as compared to more complex models. Despite their apparent lack of realism, these simple stochastic models have yielded qualitative insights into general management questions, such as the determination of minimum viable population sizes, whether demographic, genetic, or environmental fluctuations are more important to a species' survival, whether single large or several small reserves afford the least extinction risk, and how Allee effects might be manifested in stochastic populations (Leigh 1981, Wright and Hubbell 1983, Wilcox 1986, Goodman 1987, Burgman et al. 1988, Simberloff 1988, Burkey 1989, Dennis 1989a).

The usefulness of simple stochastic models, though, seemingly diminishes for specific, real situations. First, many endangered vertebrate populations are age structured and have periodic breeding seasons, so that models containing a single state variable and/or continuous time would appear unrealistic and inappropriate for quantitative predictions concerning particular species. An alternate approach preferred by some investigators has been the construction of detailed simulation models with many variables, parameters, and stochastic components (e.g., Shaffer and Samson 1985, Mode and Jacobson 1987a, b, Ferson et al. 1989). Second, stochastic models are limited in practice by the quality and quantity of data on the growth of endangered species. In this regard, the improvements offered by detailed simulation models over simple analytical models can be dubious. It is a statistical fact of life that less data allows estimation of fewer parameters, and it is quite common for components (particularly stochastic components) of simulation models to be set rather arbitrarily. Finally, even in situations where good data exist, it has not been at all clear to investigators how to interface stochastic models (or deterministic ones, for that matter) with data. Appropriate statistical methods for parameter estimation, model evaluation, and hypothesis testing for stochastic growth models would enhance our understanding of biological populations.

However, mathematical studies have suggested that simple stochastic models might serve as useful approximations for various quantities pertaining to age-structured populations experiencing stochastic fluctuations. Theoretical results concerning stochastic Leslie matrices have indicated that, under very mild assumptions, the logarithm of total population size in an age-structured population can be approximated by a simple stochastic process known as the Wiener process (or Brownian motion) with drift (Tuljapurkar and Orzack 1980, Heyde and Cohen 1985, Tuljapurkar 1989). This model represents a simple stochastic exponential growth model for population size, and as such had been proposed and analyzed in the population biology literature (Capocelli and Ricciardi 1974). Most of its statistical properties, like the distribution of the waiting time until an upper or lower fixed size is attained, had been derived decades earlier by physicists and mathematicians. Population biologists have realized that these statistical properties might prove useful for extinction calculations (Capocelli and Ricciardi 1974, Ginzburg et al. 1982, Braumann 1983b, Levinton and Ginzburg 1984, Lande and Orzack 1988). Lande and Orzack (1988), in particular, recently emphasized the potential importance of the Wiener-drift process as a general approximation for age-structured populations, and showed with computer simulations that the associated approximations for extinction-related quantities can be quite acceptable.

Furthermore, statistical inference methods for simple stochastic processes are accumulating in the mathematical statistics literature (e.g., Basawa and Prakasa-Rao 1980), and inference for the Wiener-drift process has been developed in the specific context of population biology. Braumann (1983b) and Dennis (1989b) studied the question of how to fit this stochastic exponential growth model to data. To "fit" means to estimate the two parameters (denoted μ and σ^2 in this paper) in the model in some statistically acceptable manner, given data on some growing or declining population. Braumann (1983b) derived maximum likelihood (ML) estimates for time series data with observations spaced at equal intervals. Dennis (1989b) generalized these estimates to unequally spaced intervals and showed how the problem can be transformed to a simple linear regression, making available the whole battery of linear model diagnostics, tests, and software.

These combined developments now make possible the estimation of quantities related to growth and extinction for a variety of endangered species, using relatively straightforward statistical techniques. We describe the necessary techniques in this paper, and provide illustrative analyses of data on the Whooping Crane (Grus americana), grizzly bear (Ursus arctos horribilis), Kirtland's Warbler (Dendroica kirtlandii), California Condor (Gymnogyps californianus), Puerto Rican Parrot (Amazona vittata), Palila (Loxioides bailleui), and Laysan Finch (Telespyza cantans). The quantities estimated are functions of the two parameters in the

stochastic exponential growth model and include the continuous rate of increase, the finite rate of increase, the probability of extinction, the mean time to extinction, and the projected population size. The model is easy to use and fits the example data well. The statistical methods discussed here for applying the model are best used with true census data, or with population estimates in situations where sampling variability is small compared to population variability. We conclude that the model, in conjunction with the statistical inference methods described here, is a potentially valuable tool for addressing scientific and management questions in conservation biology.

THE STOCHASTIC EXPONENTIAL GROWTH MODEL

Projection matrix

The Lewis-Leslie model (Lewis 1942, Leslie 1945) is a frequently used mathematical representation of density-independent growth of an age-structured population observed at discrete time intervals. The model can be written as

$$M(t + 1) = A(t)M(t), \qquad (1)$$

where $M(t)$ is a column vector containing elements representing numbers of individuals (usually females) in each age class at time t ($t = 0, 1, 2, \ldots$), and $A(t)$ is a square matrix containing age-specific fecundity rates (top row), age-specific survivorship rates (subdiagonal), and zeros elsewhere (see van Groenendael et al. 1988 for a recent review). The model is easily generalized to stage-structured populations by incorporating additional positive elements into the projection matrix $A(t)$ (Lefkovitch 1965). If the elements in $A(t)$ are constant, the total population size ultimately approaches exponential growth or decline, after initial age- (or stage-) structure imbalances damp out into a stable age structure. The exponential growth is represented by

$$N(t) = n_0 \lambda^t, \qquad (2)$$

where $N(t)$ is the total population size [summed elements of the vector $M(t)$, $n_0 = N(0)$ is the initial population size, and λ is the dominant eigenvalue of the projection matrix (finite rate of increase).

However, the elements of a realistic projection matrix should fluctuate with time, since fecundity, survivorship, or stage transition rates are seldom constant in nature. An alternative modeling approach is to assume that the elements of $A(t)$ change with time in the form of a (multivariate) stationary time series. Tuljapurkar (1989) has given a comprehensive review of the demographic theory of populations governed by such dynamics. This modeling assumption is mathematically broad enough to include many real situations and is fundamental to the analysis methods we describe in this paper. Note that this assumption excludes populations experiencing nonstationary fluctuations in de-

mographic rates, such as decreasing survival or reproduction rates due to diminishing habitat.

This stochastic formulation seems to have the minimum level of biological detail necessary for describing a vertebrate population. By contrast, a deterministic, single-state variable model such as Eq. 2 would not likely provide much useful information concerning survival or extinction of an endangered species.

A stochastic single-state variable model, though, can adequately approximate the statistical properties of the fluctuations in total population size resulting from the stochastic projection matrix. Results of Tuljapurkar and Orzack (1980) and Heyde and Cohen (1985), based on central limit theorems, state that the quantity $X(t) = \log N(t)$ will have, as t becomes large, an approximate normal distribution with a mean of $x_0 + \mu t$ and a variance of $\sigma^2 t$ [written $X(t) \sim \text{normal}(x_0 + \mu t, \sigma^2 t)$, where "$\sim$" means "is distributed as," and the dot indicates that the distribution is approximate], where $x_0 = \log n_0$. The approximation can be improved by adjusting n_0 for initial age structure imbalances (see Lande and Orzack 1988), but we confine ourselves in this paper to estimation techniques that do not require detailed knowledge of age structure. Here, μ is a real-valued constant, and σ^2 is a positive, real-valued constant. These parameters depend on properties of the underlying stochastic projection matrix. If the matrices $A(1), A(2), \ldots$, are serially uncorrelated (for example, each year or time period, elements of the projection matrix are drawn from a multivariate distribution, independent of previous years), then

$$\mu \approx \log \lambda - (\sigma^2/2) \qquad (3)$$

and

$$\sigma^2 \approx \lambda^{-2} \delta' C \delta, \qquad (4)$$

where λ is now the dominant eigenvalue of the average projection matrix $A^*\{= E[A(t)]\}$, C is the variance–covariance matrix of the multivariate distribution from which the elements of $A(t)$ arise, and δ is a column vector containing partial derivatives of λ with respect to each element of A^* (Tuljapurkar 1982b).

While estimating the multitudinous quantities in a projection matrix with any useful degree of precision can be exceedingly difficult, estimating μ and σ^2 is possible with just a single time series of observations on total population size. Furthermore, various quantities related to extinction are functions of μ and σ^2 and are straightforwardly estimated.

Ease of estimation arises from the fact that the approximate normal distribution of $X(t)$ is identical to the distribution of a Wiener process with drift (e.g., Goel and Richter-Dyn 1974). The Wiener-drift model is a simple type of continuous-time, continuous-state, Markov stochastic process known as a diffusion process. Taking $X(t)$ to be a Wiener-drift process, strictly speaking, imposes an additional layer of approximation on top of the results of Tuljapurkar and Orzack

(1980) and Heyde and Cohen (1985), in that the Wiener-drift process has the above-mentioned normal distribution for small values of t as well as large. Lande and Orzack (1988), though, showed with computer simulations that extinction probabilities under various hypothetical life histories were accurately predicted with the Wiener-drift approximation, provided the fluctuations in the projection matrix elements were small or moderate. We discuss diagnostic procedures later in this paper to evaluate the adequacy of the Wiener-drift model for a given data set (see *Model evaluation* section).

Diffusion approximation

We assume that the natural logarithm of total population size, $X(t) = \log N(t)$, is adequately approximated by a Wiener process with drift. This process has been extensively studied; Goel and Richter-Dyn (1974), Ricciardi (1977), and Karlin and Taylor (1981) provide lucid expositions of it and other diffusion processes. The constant μ is known as the infinitesimal mean of the process, since $\mu \Delta t$ is the (approximate) average amount of change in the process over a tiny time interval Δt. The constant σ^2 likewise is known as the infinitesimal variance. The process has a transition probability density function (pdf) corresponding to a normal$(x_0 + \mu t, \sigma^2 t)$ distribution:

$$p_X(x, t \mid x_0) = (2\pi\sigma^2 t)^{-\frac{1}{2}}\exp[-(x - x_0 - \mu t)^2/(2\sigma^2 t)],$$
$$-\infty < x < \infty. \quad (5)$$

The probability that $X(t)$ is between a and b at time t, given that the process starts at x_0, is the corresponding area under the pdf. This probability can be evaluated with the standard normal cumulative distribution function (cdf), $\Phi(\cdot)$:

$$\Pr[a < X(t) \leq b] = \int_a^b p_X(x, t \mid x_0)\, dx$$
$$= \Phi\left(\frac{b - x_0 - \mu t}{\sqrt{\sigma^2 t}}\right)$$
$$- \Phi\left(\frac{a - x_0 - \mu t}{\sqrt{\sigma^2 t}}\right), \quad (6)$$

where

$$\Phi(z) = \int_{-\infty}^z (2\pi)^{-\frac{1}{2}}\exp(-y^2/2)\, dy. \quad (7)$$

The untransformed total population size, $N(t) = \exp[X(t)]$, is also a diffusion process. The transition pdf for $N(t)$ is that of a lognormal distribution:

$$p_N(n, t \mid n_0)$$
$$= n^{-1}(\sigma^2 t 2\pi)^{-\frac{1}{2}}\exp[-(\log n - \log n_0 - \mu t)^2/(2\sigma^2 t)],$$
$$0 < n < \infty. \quad (8)$$

The mean (or expected) population size, given that the process starts at n_0, is

$$E[N(t)] \equiv \psi(t; n_0, \mu, \sigma^2)$$
$$= n_0\exp\{[\mu + (\sigma^2/2)]t\}. \quad (9)$$

For the case of serially uncorrelated projection matrices, the mean population size is identical to the deterministic exponential growth model (Eq. 2), with λ given by Eq. 3. The mean of $X(t)$ by contrast does not depend on σ^2:

$$E[X(t)] \equiv \gamma(t; x_0, \mu) = x_0 + \mu t. \quad (10)$$

The geometric mean of $N(t)$ is defined by $\exp\{E[X(t)]\}$:

$$\exp\{E[\log N(t)]\} \equiv \beta(t; n_0, \mu) = n_0\exp(\mu t). \quad (11)$$

The geometric mean is also the median of the lognormal transition pdf (Eq. 8), which is a general property of the lognormal distribution (see Dennis and Patil 1988). Analytical (Tuljapurkar 1982*a*) and simulation (Slade and Levenson 1982, Nordheim et al. 1989) studies suggest that the geometric mean of $N(t)$ better characterizes the behavior of the process than does the mean (Eq. 9), due to the extreme positive skewness of the lognormal transition pdf. Other statistical properties of this process are catalogued by Dennis and Patil (1988).

Any diffusion process in general has an alternate mathematical representation as a stochastic differential equation (SDE) (e.g., Karlin and Taylor 1981). Of particular interest here is that the process $N(t)$ is the solution to an SDE version of the exponential growth model given by

$$dN(t) = rN(t)\, dt + \sigma N(t)\, dW(t), \quad (12)$$

where r is a real-valued constant and $dW(t) \sim$ normal$(0, dt)$. The differential $dN(t)$ is defined mathematically in terms of an Ito stochastic integral (for example, Soong 1973); the constants μ and σ^2 in the transition pdf (Eq. 8) for $N(t)$ are related to r by

$$r = \mu + (\sigma^2/2). \quad (13)$$

For the case of serially uncorrelated projection matrices (see Eq. 3), $r \approx \log \lambda$.

Previous studies of extinction probabilities (Capocelli and Ricciardi 1974) and parameter estimation (Braumann 1983*b*) for the exponential growth SDE (Eq. 12) used a Stratonovich stochastic integral to define $dN(t)$; those results should not be used if the diffusion process $N(t)$ is intended as an approximation for an underlying age-structured, discrete-time system. The Stratonovich formulation instead more appropriately represents a system in which the state variable is fundamentally a continuous function of time (e.g., biomass). Turelli (1977), Capocelli and Ricciardi (1979), and Braumann (1983*a*) provide insights into the two ways of interpreting the SDE (Eq. 12) as an approximation for some underlying process, and Dennis and Patil (1988) list formulas for transforming Ito-based results to Stratonovich-based results and vice versa.

Extinction properties

Under the continuing unpredictable fluctuations of the Wiener-drift model, $X(t)$ could possibly cross any lower threshold size, x_e, starting from x_0. This event corresponds to the population size, $N(t)$, attaining a lower threshold size, $n_e = \exp(x_e)$, starting from n_0. If $n_e = 1$ (or $x_e = 0$), the event obviously represents the extinction of a closed, sexually reproducing population. Management efforts to promote survival of an endangered species might naturally hinge upon a different threshold size. Some fixed population size, $n_e > 1$, could be regarded as a policy threshold, or as a safety cushion to avoid the possibilities of Allee effects (e.g., Dennis 1989a), skewed sex ratios, or inbreeding. We use the term "extinction" in this paper to refer broadly to the attainment of some prespecified lower threshold, representing, if not the demise of the species, the demise of some management regime. The term "quasiextinction" has also been used in this context (Ginzburg et al. 1982). Let x_d represent the distance on the logarithmic scale from an initial population size to a lower threshold population size:

$$x_d = x_0 - x_e = \log(n_0/n_e). \qquad (14)$$

As explained in the context of extinction by numerous authors (Capocelli and Ricciardi 1974, Ricciardi 1977, Tuljapurkar and Orzack 1980, Ginzburg et al. 1982, Lande and Orzack 1988), the probability $\pi(x_d, \mu, \sigma^2)$ that the process will ever attain the threshold is

$$\pi(x_d, \mu, \sigma^2) = \begin{cases} 1, \mu \le 0; \\ \exp(-2\mu x_d/\sigma^2), \mu > 0. \end{cases} \qquad (15)$$

Given that the threshold is attained (i.e., conditioning on all sample paths of the process that reach the threshold), the amount of time, T, elapsing before the threshold is first reached is a positive, real-valued random variable with a continuous probability distribution. The cdf of the distribution can be written in terms of a standard normal cdf:

$$\Pr[T \le t] \equiv G(t; x_d, \mu, \sigma^2)$$

$$= \Phi\left(\frac{-x_d + |\mu|t}{\sigma\sqrt{t}}\right)$$

$$+ \exp(2x_d|\mu|/\sigma^2)\Phi\left(\frac{-x_d + |\mu|t}{\sigma\sqrt{t}}\right),$$

$$0 < t < \infty. \quad (16)$$

The pdf of the distribution is the derivative of $G(t; x_d, \mu, \sigma^2)$ with respect to t:

$$g(t; x_d, \mu, \sigma^2)$$
$$= x_d(2\pi\sigma^2t^3)^{-1/2}\exp[-(x_d - |\mu|t)^2/(2\sigma^2t)]. \quad (17)$$

This distribution, known as the inverse Gaussian distribution (a misnomer: it is not the distribution of the reciprocal of a normal random variable), has been extensively studied (see Folks and Chhikara 1978).
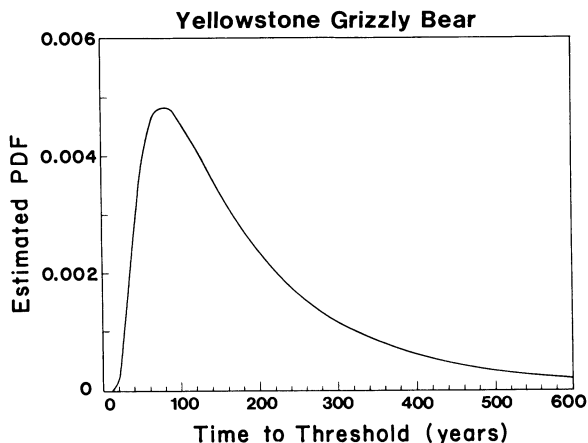

FIG. 1. Probability density function (PDF) of the inverse Gaussian distribution, plotted using maximum likelihood estimates of parameters μ and σ^2 for the Yellowstone National Park population of grizzly bears. The distribution is that of the time required for 47 female bears to decline to 10 bears.

Schrödinger (1915) originally obtained it as the first-passage time distribution for the Wiener-drift process for the case $\mu < 0$; Tweedie (1957a, b) derived many of its statistical properties. Whitmore (1978) obtained the inverse Gaussian explicitly as the conditional first-passage time distribution for the case $\mu > 0$. Whitmore and Seshadri (1987) and Lande and Orzack (1988) have given intuitive derivations of the first-passage time result.

We note that the probability distribution of the time to attain an upper threshold, given it is attained, is also the inverse Gaussian distribution (Eqs. 16 and 17), except with x_d representing $\log(n_u/n_0)$, where n_u is the upper threshold. The probability of ever attaining n_u is given by $\pi(x_d, -\mu, \sigma^2)$, that is, by (Eq. 15) evaluated at $\log(n_u/n_0)$, $-\mu$, and σ^2.

Various quantities pertaining to this distribution are of potential interest in conservation biology. The mean time until the threshold population size is reached is the expected value of T:

$$E[T] \equiv \theta(x_d, \mu) = x_d/|\mu|. \qquad (18)$$

The variance of T is

$$\text{Var}(T) = x_d\sigma^2/|\mu|^3. \qquad (19)$$

The distribution is positively skewed and has a heavy right tail (an example shape is portrayed in Fig. 1), a fact that has implications for species preservation efforts (see *Examples* and *Discussion* sections). Percentiles and modes are quantities that help to characterize such skewed distributions. The $100 \cdot p$th percentile, $\xi_p(x_d, \mu, \sigma^2)$, is defined as the root of

$$G(\xi_p; x_d, \mu, \sigma^2) = p. \qquad (20)$$

For example, the 50th percentile of the distribution, or the median, represents the fixed time at which the probability of hitting the threshold before that time is 0.5. The mode of the distribution is the most likely time of hitting the threshold, and is defined as the value of t maximizing the pdf (Eq. 17):

$$t^*(x_d, \mu, \sigma^2) = \frac{x_d}{|\mu|}\left[\left(1 + \frac{9}{4\nu^2}\right)^{\!\!{1/2}} - \frac{3}{2\nu}\right], \quad (21)$$

where $\nu = x_d|\mu|/\sigma^2$. Note that the mode is the product of the mean and a quantity (in square brackets) between 0 and 1.

The notation used above emphasizes that these quantities are functions of the parameters x_d, μ, and σ^2. The quantity x_d, the (log-scale) distance to the threshold population, is selected by the investigator. The parameters μ and σ^2, however, are typically unknown and must be estimated from data.

PARAMETER ESTIMATION

Maximum likelihood estimates

Estimates of μ and σ^2 can be obtained by observing a population at times $0, t_1, t_2, \ldots, t_q$. The recorded observations of population size will be denoted $n(0) = n_0$, $n(t_1) = n_1, \ldots, n(t_q) = n_q$, and the time intervals (not necessarily equal) between observations denoted $t_1 - 0 = \tau_1$, $t_2 - t_1 = \tau_2, \ldots, t_q - t_{q-1} = \tau_q$. For a population with a yearly breeding cycle, these observations should be spaced at least 1 yr apart and taken at the same time of year. A recommended way of fitting the stochastic exponential growth model to such data is maximum likelihood (ML) estimation.

ML estimates of the two parameters are easy to calculate and have many desirable statistical properties. The likelihood function $l(\mu, \sigma^2)$ is defined as the joint pdf for $N(t_1), N(t_2), \ldots, N(t_q)$, given $N(0) = n_0$, evaluated at the observations $n_1, n_2, \ldots, n_q$. To obtain the likelihood function, we note that $N(t)$ is a diffusion process with stationary transition probabilities. This means that the pdf of n_i given n_{i-1} depends on τ_i (the time interval since n_{i-1}) but not on t_{i-1}. Also, the diffusion process is a Markov process, meaning that the pdf of n_i given n_{i-1} does not depend on the earlier observations $n_0, \ldots, n_{i-2}$. Thus, $p_N(n_i, \tau_i | n_{i-1})$, which is the lognormal transition pdf (Eq. 8) evaluated at n_i, τ_i, and n_{i-1}, represents the likelihood of the system moving to n_i from n_{i-1} in a time interval of τ_i. The likelihood function is then the product of transition pdfs:

$$l(\mu, \sigma^2)$$
$$= p_N(n_1, \tau_1 | n_0)p_N(n_2, \tau_2 | n_1)\cdots p_N(n_q, \tau_q | n_{q-1}). \quad (22)$$

The ML estimates, $\hat{\mu}$ and $\hat{\sigma}^2$, are the parameter values jointly maximizing $l(\mu, \sigma^2)$ or, equivalently log $l(\mu, \sigma^2)$, where

$$\log l(\mu, \sigma^2) = -\sum_{i=1}^{q} \log[n_i(2\tau_i\pi)^{1/2}]$$
$$- (q/2)\log \sigma^2 - [1/(2\sigma^2)]$$
$$\cdot \sum_{i=1}^{q} (1/\tau_i)[\log(n_i/n_{i-1}) - \mu\tau_i]^2. \quad (23)$$

It is straightforward to set partial derivatives of log $l(\mu,\sigma^2)$ with respect to μ and σ^2 equal to zero and solve for the ML estimates:

$$\hat{\mu} = \left[\sum_{i=1}^{q} \log(n_i/n_{i-1})\right]\Big/ \sum_{i=1}^{q} \tau_i$$
$$= [\log(n_q/n_0)]/t_q; \quad (24)$$

$$\hat{\sigma}^2 = (1/q) \sum_{i=1}^{q} (1/\tau_i)[\log(n_i/n_{i-1}) - \hat{\mu}\tau_i]^2. \quad (25)$$

We mention that the ML estimate (Eq. 25) of σ^2 is biased (e.g., Dennis 1989b); an unbiased estimate is

$$\tilde{\sigma}^2 = q\hat{\sigma}^2/(q - 1). \quad (26)$$

The difference between $\hat{\sigma}^2$ and $\tilde{\sigma}^2$ is negligible when q is large. Often q will not be very large in data on endangered populations, though.

Linear regression approach

These identical ML estimates of μ and σ^2 can also be calculated by a regression approach. This approach offers the following practical advantages. First, information about the statistical distributions of the ML estimates is easily obtained (e.g., confidence intervals). Second, a battery of diagnostic procedures for linear regression models becomes available for evaluating the diffusion model. Finally, the analyses can be accomplished with most of the standard computer packages for linear regression.

The approach involves transforming the observations so that a normal linear model applies. Let $W_i = \log[N(t_i)/N(t_{i-1})] = X(t_i) - X(t_{i-1})$, so that W_i represents the change in $X(t)$ (the Wiener-drift process) between times t_{i-1} and t_i. Thus, the variables $W_1, W_2, \ldots, W_q$ are increments of a Wiener-drift process, and are therefore normal, independent, and stationary (e.g., Ricciardi 1977). In fact, if $\mathbf{w} = [W_1, W_2, \ldots, W_q]'$ and $\boldsymbol{\tau} = [\tau_1, \tau_2, \ldots, \tau_q]'$ are defined as column vectors, the distribution of $\mathbf{w}$ becomes a multivariate normal, with mean $\mu\boldsymbol{\tau}$ and variance–covariance matrix $\sigma^2\mathbf{V}$:

$$\mathbf{w} \sim \text{normal}(\mu\boldsymbol{\tau}, \sigma^2\mathbf{V}). \quad (27)$$

Here $\mathbf{V} = \text{diag}(\boldsymbol{\tau})$ is a matrix with the elements of $\boldsymbol{\tau}$ on the main diagonal and zeros elsewhere. Let $\mathbf{G} = \text{diag}(\sqrt{\tau_1}, \ldots, \sqrt{\tau_q})$, that is, $\mathbf{V} = \mathbf{G}'\mathbf{G}$. A transformation of $\mathbf{w}$ produces an ordinary normal linear model (e.g., Graybill 1976:207):

$$\mathbf{Y} = \mathbf{G}^{-1}\mathbf{w} \sim \text{normal}(\mu\mathbf{D}, \sigma^2\mathbf{I}), \quad (28)$$

where $\mathbf{D} = [\sqrt{\tau_1}, \ldots, \sqrt{\tau_q}]'$ and $\mathbf{I}$ is the $q \times q$ identity

matrix. This is a model for a simple linear regression without intercept.

In practice, the data are transformed by $y_i = [\log(n_i/n_{i-1})]/\sqrt{\tau_i}$, $i = 1, \ldots, q$. The regression approach uses $y_1, y_2, \ldots, y_q$ as values of the "dependent variable," $\sqrt{\tau_1}, \ldots, \sqrt{\tau_q}$ as values of the "independent variable," and a linear regression without intercept is performed. The formula (Eq. 24) for $\hat{\mu}$ is recognized as the slope parameter estimate, and $\tilde{\sigma}^2$ (Eq. 26) is the (unbiased) estimate of the error variance parameter. The ML estimate of σ^2 is $\hat{\sigma}^2$ (Eq. 25).

The usual linear model theory yields the distributions of $\hat{\mu}$, $\hat{\sigma}^2$, and $\tilde{\sigma}^2$ (Graybill 1976):

$$\hat{\mu} \sim \text{normal}(\mu, \sigma^2/t_q), \qquad (29)$$

$$q\hat{\sigma}^2/\sigma^2 = (q - 1)\tilde{\sigma}^2/\sigma^2 \sim \text{chi-square}(q - 1). \qquad (30)$$

Also, $\hat{\mu}$ is independent of $\hat{\sigma}^2$ or $\tilde{\sigma}^2$. An estimate of the standard error of $\hat{\mu}$ is $(\tilde{\sigma}^2/t_q)^{1/2}$, and a $100(1 - \alpha)\%$ confidence interval for μ is given by

$$(\hat{\mu} - t_{\alpha/2,q-1}\sqrt{\tilde{\sigma}^2/t_q}, \ \hat{\mu} + t_{\alpha/2,q-1}\sqrt{\tilde{\sigma}^2/t_q}). \qquad (31)$$

Here $\Pr[|T_{q-1}| \leq t_{\alpha/2,q-1}] = 1 - \alpha$, where T_{q-1} has a Student's t distribution with $q - 1$ degrees of freedom. Confidence intervals for σ^2 can be calculated with either the ML or the unbiased estimate. For example, a $100(1 - \alpha)\%$ confidence interval for σ^2 based on the ML estimate is

$$(q\hat{\sigma}^2/\chi^2_{\alpha_1,q-1}, \ q\hat{\sigma}^2/\chi^2_{1-\alpha_2,q-1}), \qquad (32)$$

where $\alpha_1 + \alpha_2 = \alpha$, $0 < \alpha_1, \alpha_2$, and $\Pr[X_{q-1} \leq \chi^2_{\alpha,q-1}] = 1 - \alpha$, where X_{q-1} has a chi-square distribution with $q - 1$ degrees of freedom. The confidence interval using the unbiased estimate would simply substitute $(q - 1)\tilde{\sigma}^2$ for $q\hat{\sigma}^2$ in Eq. 32.

MODEL EVALUATION

The equivalence of the Wiener-drift model for $X(t)$ and the linear regression model for the transformed increment variables Y (see Eq. 28) is extremely useful for assessing the adequacy of the diffusion approximation. Numerous diagnostic procedures are available for evaluating the adequacy of linear models; Chatterjee and Hadi (1988), Cook and Weisburg (1982), Belsley et al. (1980), and Draper and Smith (1981) are excellent references. Furthermore, changes in the Wiener-drift model, before and after a fixed time, can be detected using regression methods. The relevance of these procedures to evaluating the diffusion model is discussed.

Evaluation of model assumptions

Our assumption of a Wiener-drift model for $X(t)$ corresponds to the following regression model for the transformed increment data (from Eq. 28):

$$Y = \mu D + \epsilon, \qquad (33)$$

where the $q \times 1$ vector of errors $\epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_q)'$;

the errors are assumed to be independent normal random variables with common mean 0 and variance σ^2. Thus, a violation of the diffusion model assumptions can be detected by evaluating the assumptions of the regression analysis.

Notice that the transition is the fundamental observation in the likelihood function (Eq. 22); the population observations $n_0, n_1, \ldots, n_q$ appear only in the context of transitions from one to the next.

Graphical and analytical methods for checking the regression model assumptions are invariably based on the residuals or transformations of the residuals, where the jth residual is the difference between the jth observation and its predicted value: $e_j = y_j - \hat{\mu}\sqrt{\tau_j}$ (e.g., Draper and Smith 1981, Cook and Weisburg 1982). Although the residuals are subject to some restrictions (e.g., correlations among the e_j may be nonzero), they can be loosely interpreted as the observed errors if the model is correct. Hence, they should exhibit behaviors that tend to confirm the model assumptions or at least that do not contradict these assumptions (Draper and Smith 1981).

Residuals can be used to check the independent increments property, an assumption imposed on $X(t)$ via the Wiener-drift approximation. This assumption is crucial to the analyses presented here. If the assumption is reasonable, then serial autocorrelations among the increment variables W_j (see Eq. 27) are negligible and hence the limiting distributional results of Tuljapurkar and Orzack (1980) and Heyde and Cohen (1985) can be applied to the transitions; that is, the W_j are approximately independent normal random variables. Correlations among the transitions suggest that the Wiener-drift process for $X(t)$ is inappropriate. In this case, while the least squares estimates of μ and σ^2 are statistically consistent (i.e., the estimates get "closer" to the true parameter values as the sample size increases), they may not be the best estimates. We refer to Heyde and Cohen (1985) for more appropriate estimators when the errors are correlated.

We note that even with serially uncorrelated environmental fluctuations, the W_j in most age-structured populations have a theoretical serial correlation. The cause of this correlation is the time lag inherent in age-structured survival and reproduction; a big pulse of reproduction, for instance, leads to another pulse some years later when the "baby boom" reaches reproductive maturity. The theoretical Wiener-drift approximation strictly applies only for time intervals encompassing many generations. Use of a Wiener-drift likelihood (Eq. 27) to model the W_j will result, in the case of serially uncorrelated environmental fluctuations, in an underestimate of the infinitesimal variance of the theoretical Wiener-drift approximation.

There is as yet no entirely satisfactory way to correct for this bias with a single time series, without additional knowledge of age-specific demographic parameters. Methods discussed by Heyde and Cohen (1985) and

Lande and Orzack (1988) involve estimating high order autocovariances or pooling transitions into a few long-term transitions, with the price of high standard errors for the estimates of σ^2. The situation warrants further investigation. In the meantime, use of the Wiener-drift model for a population time series should rely heavily on diagnostic techniques, such as discussed here, in order to minimize potential bias.

A popular test for detecting first-order autocorrelation among the errors of a regression model is the Durbin-Watson test. The test statistic for testing the null hypothesis of uncorrelated errors is

$$d = \sum_{j=2}^{q} (e_j - e_{j-1})^2 \Big/ \sum_{j=1}^{q} e_j^2. \qquad (34)$$

Draper and Smith (1981) provide tables of upper and lower critical values of d for a number of significance levels. This test is available in many statistical computing packages and therefore can be easily incorporated in the analysis. Higher order autocorrelations can be detected by subjecting the residuals to standard time-series analyses (e.g., Pankratz 1983), provided enough data are available.

Other analytical, as well as graphical methods are available for checking the remaining model assumptions (e.g., normality, constant variance) and we refer the reader to Draper and Smith (1981) and Cook and Weisburg (1982) for a discussion of these standard procedures. It is worth noting that if the population observations $n_0, n_1, \ldots, n_q$ are taken at equally spaced time points, so that the τ_j are equal, many of these graphical methods are uninformative.

Sensitivity analysis

In recent years, numerous statistical measures have been developed for detecting unusual (e.g., outliers) or highly influential observations (Belsley et al. 1980, Chatterjee and Hadi 1988). As pointed out by Belsley et al. (1980:3), such transitions are not necessarily "bad" data points; rather, they may contain some of the most interesting sample information. However, they may also reflect an unknown recording error or they may have resulted from circumstances different from those common to the remaining data (e.g., a population reduction due to an unusual catastrophic storm or removal of population members into captivity by managers). Since these observations can have a substantial effect on the parameter estimates (that is, the estimates are extremely sensitive to them), we recommend screening for them in the analysis; many statistical software packages will calculate these measures.

An observation is an outlier if it is not successfully accommodated by the fitted regression model. Residuals or transformed residuals corresponding to outliers are large compared to those of other observations in the data set. However, a small residual does not necessarily imply that the corresponding observation is a typical one; the method of least squares avoids large

residuals in fitting the model and thus may accommodate an atypical transition at the expense of the other data points. These transitions are commonly referred to as influential observations since they excessively influence the parameter estimates as compared to the remaining data. Note that a transition may be judged as an outlier, an influential observation, or both.

Outliers are identified via statistical measures based on residuals and can be detected using formal testing procedures or informal comparisons of relative magnitude. One measure, which Chatterjee and Hadi (1988: 74) call the internally studentized residual, is the jth residual divided by its estimated standard deviation:

$$I_j = \frac{e_j}{\sqrt{\tilde{\sigma}^2 (1 - h_{jj})}}, \qquad (35)$$

where $h_{jj} = \tau_j/t_q$ is the jth diagonal element of the matrix $D(D'D)^{-1}D'$ (see Eq. 28). While the I_j are not independent (since the residuals are not independent), a formal test is available for detecting the presence of a single outlier. If I_{max} denotes the maximum of the values of $|I_j|$, then approximate critical values for I_{max}, at the α significance level, are given by

$$c_\alpha = \sqrt{\frac{(q-1) f_{\alpha/q,1,q-1}}{q - 2 + f_{\alpha/q,1,q-1}}}, \qquad (36)$$

where $f_{\alpha/q,1,q-1}$ is the $100[1 - (\alpha/q)]$th percentile of an F distribution with 1 and $q - 1$ degrees of freedom. Thus, at the α level of significance, I_{max} is an outlier if $I_{max} > c_\alpha$.

The externally studentized residual (Chatterjee and Hadi 1988:74) is defined as

$$E_j = \frac{e_j}{\sqrt{\tilde{\sigma}^2_{(j)} (1 - h_{jj})}}, \qquad (37)$$

where $\tilde{\sigma}^2_{(j)}$ is the (unbiased) estimate of σ^2 when the jth transition is deleted from the analysis (see Eq. 42 below). This measure for detecting outliers has some advantages over I_j. By excluding the jth transition in estimating σ^2, $\tilde{\sigma}^2_{(j)}$ ignores gross errors in the jth observation. Also, E_j tends to reflect large deviations more dramatically than I_j. In addition, E_j has a Student's t distribution with $q - 2$ degrees of freedom, which suggests that a transition for which $|E_j| > 2$ and E_j is large in magnitude compared to those of the remaining transitions should be investigated as a possible outlier.

Measures for detecting influential observations are commonly based on the omission approach, in that they measure changes in the parameter estimates or predicted values when the jth data point is excluded from the analysis. Cook's distance, C_j (Chatterjee and Hadi 1988:117), measures the change in the slope estimate, $\hat{\mu}$, and can be expressed as a function of the internally studentized residual:

$$C_j = I_j^2 h_{jj}/(1 - h_{jj}). \qquad (38)$$

Large values of C_j indicate that the jth transition is influential. Although C_j is not a true F random variable, comparing C_j to the probability points of an F distribution with 1 and $q - 1$ degrees of freedom provides descriptive or qualitative levels of significance (Chatterjee and Hadi 1988).

Welsch and Kuh's distance, WK_j, or the $DFFITS_j$ statistic (Chatterjee and Hadi 1988:120), measures the influence of the jth observation on the predicted value $\hat{\mu}\sqrt{\tau_j}$:

$$WK_j = |E_j| \sqrt{\frac{h_{jj}}{1 - h_{jj}}} . \qquad (39)$$

Again, a large value of this statistic indicates a potentially influential transition. While WK_j does not have a Student's t distribution with $q - 2$ degrees of freedom, it is a t-like statistic. This suggests that data points for which $WK_j > 2$ should be regarded as influential observations.

In some regression studies h_{jj} alone is used as an influence measure. A point is considered influential if h_{jj} exceeds some specified constant. We point out that this approach leads to some peculiarities for regression through the origin. Since $h_{jj} = \tau_j/t_q$, only transitions with longer time intervals can be influential, but never shorter time intervals. The other influence measures discussed in this subsection depend on the y_j values as well as the h_{jj} values and hence avoid this undesirable property.

Once unusual or highly influential transitions have been identified, they must be investigated to determine if they are in error or the result of some catastrophic event. If so, we recommend deleting them from the analysis. When the jth transition is deleted, the likelihood function (Eq. 22) has the corresponding transition pdf excluded from the product, and the log-likelihood (Eq. 23) has the jth term excluded from the sum. Eqs. 24 and 25 for the ML estimates then become altered to account for the missing transition:

$$\hat{\mu} = \left\{ \sum_{\substack{i=1 \\ i \neq j}}^{q} \log(n_i/n_{i-1}) \right\} \bigg/ \left\{ \sum_{\substack{i=1 \\ i \neq j}}^{q} \tau_i \right\}; \qquad (40)$$

$$\hat{\sigma}^2 = (1/p) \sum_{\substack{i=1 \\ i \neq j}}^{q} (1/\tau_i)[\log(n_i/n_{i-1}) - \hat{\mu}\tau_i]^2. \qquad (41)$$

Here p is the number of transitions included in the estimates ($= q - 1$ when just one transition is deleted). Additional transitions can be deleted from the formulas in the same way. If transitions are deleted from the ML estimate of σ^2, then the unbiased estimate is

$$\tilde{\sigma}^2 = p\hat{\sigma}^2/(p - 1). \qquad (42)$$

Note that $\tilde{\sigma}^2 = \tilde{\sigma}^2_{(j)}$ if the jth observation alone is excluded. The estimates, $\hat{\mu}$ and $\tilde{\sigma}^2$ can easily be obtained by omitting the appropriate pairs $(y_j, \sqrt{\tau_j})$ from the regression analysis. With transitions deleted, the distribution Eqs. 29–32 for the parameter estimates would

have $\Sigma \tau_i$ in place of t_q, where the sum is over all transitions included in the analysis, and p in place of q.

Change in parameters

A permanent change in the infinitesimal mean growth rate may result from a sustained change in the population's environment or from a change in conservation efforts on the part of managers. Suppose the change in conditions is known to have occurred after the jth transition. A corresponding change in μ (without a change in σ^2) is incorporated in the diffusion model by expanding the regression model (Eq. 33):

$$Y = D\mu + \epsilon, \qquad (43)$$

where $\mu = (\mu_1, \mu_2)'$ and D is a $q \times 2$ matrix with first column $(\sqrt{\tau_1}, \ldots, \sqrt{\tau_j}, 0, \ldots, 0)'$ and second column $(0, \ldots, 0, \sqrt{\tau_{j+1}}, \ldots, \sqrt{\tau_q})'$. Thus, μ_1 denotes the slope parameter for the first j observations while the remaining $q - j$ transitions have slope μ_2. As before, ϵ is the $q \times 1$ vector of normal random variables, with mean 0 and common variance σ^2. The ML estimates of μ_1 and μ_2 are

$$\hat{\mu}_1 = [\log(n_j/n_0)] \bigg/ \sum_{i=1}^{j} \tau_i, \qquad (44)$$

$$\hat{\mu}_2 = [\log(n_q/n_j)] \bigg/ \sum_{i=j+1}^{q} \tau_i, \qquad (45)$$

with distributions

$$\hat{\mu}_1 \sim \text{normal}\left(\mu_1, \sigma^2 \bigg/ \sum_{i=1}^{j} \tau_i\right), \qquad (46)$$

and

$$\hat{\mu}_2 \sim \text{normal}\left(\mu_2, \sigma^2 \bigg/ \sum_{i=j+1}^{q} \tau_i\right). \qquad (47)$$

The unbiased estimate of σ^2 is

$$\tilde{\sigma}^2 = (1/(q - 2))$$
$$\cdot \left\{ \sum_{i=1}^{j} (1/\tau_i)[\log(n_i/n_{i-1}) - \hat{\mu}_1\tau_i]^2 \right.$$
$$\left. + \sum_{i=j+1}^{q} (1/\tau_i)[\log(n_i/n_{i-1}) - \hat{\mu}_2\tau_i]^2 \right\}, \qquad (48)$$

while the ML estimate is

$$\hat{\sigma}^2 = (q - 2)\tilde{\sigma}^2/q. \qquad (49)$$

Also,

$$q\hat{\sigma}^2/\sigma^2 = (q - 2)\tilde{\sigma}^2/\sigma^2 \sim \text{chi-square}(q - 2). \quad (50)$$

Hence, confidence intervals for the model parameters can be readily obtained (e.g., see Eqs. 31 and 32).

A formal test is available for testing whether the change in the infinitesimal growth rate is significant. The statistic

$$T_{q-2} = (\hat{\mu}_1 - \hat{\mu}_2)/\sqrt{\tilde{\sigma}^2[(1/j) + 1/(q - j)]} \quad (51)$$

has a Student's t distribution with $q - 2$ degrees of freedom, under the null hypothesis that $\mu_1 - \mu_2 = 0$. Note that Eq. 51 is simply the two-sample t statistic (e.g., Neter et al. 1985:13).

A sustained alteration in a population's environment or management policies may also result in a permanent change in the infinitesimal variance. The model (Eq. 43) can be used to include a change in σ^2 after the jth transition by assuming that the first j errors in ϵ are normal random variables with common mean 0 and variance σ_1^2 while the remaining $q - j$ errors are also normal with mean 0 but have common variance σ_2^2. In practice, this amounts to fitting two regression models, one to the first j transitions and another to the remaining data. Allowing a corresponding change in the mean provides the best estimates of the variance parameters, since the least squares estimate of σ^2 depends on μ (see Eqs. 52 and 53). The ML estimates of $\hat{\mu}_1$ and $\hat{\mu}_2$ are Eqs. 44 and 45. Their distributions are Eqs. 46 and 47, respectively, except with separate variance σ_1^2 and σ_2^2 substituted in place of σ^2. The unbiased estimates of the variances are

$$\tilde{\sigma}_1^2 = [1/(j - 1)] \sum_{i=1}^{j} (1/\tau_i)$$
$$\cdot [\log(n_i/n_{i-1}) - \hat{\mu}_1 \tau_i]^2, \qquad (52)$$

$$\tilde{\sigma}_2^2 = [1/(q - j - 1)] \sum_{i=j+1}^{q} (1/\tau_i)$$
$$\cdot [\log(n_i/n_{i-1}) - \hat{\mu}_2 \tau_i]^2, \qquad (53)$$

with independent distributions

$$(j - 1)\tilde{\sigma}_1^2/\sigma_1^2 \sim \text{chi-square}(j - 1), \qquad (54)$$

$$(q - j - 1)\tilde{\sigma}_2^2/\sigma_2^2 \sim \text{chi-square}(q - j - 1). \qquad (55)$$

Since the chi-square random variables Eqs. 54 and 55 are independent, the ratio of the estimated variances given by

$$F_{j-1,q-j-1} = \tilde{\sigma}_1^2/\tilde{\sigma}_2^2 \qquad (56)$$

has an F distribution with $j - 1$ and $q - j - 1$ degrees of freedom, under the null hypothesis that $\sigma_1^2 = \sigma_2^2$ (e.g., Neter et al. 1985:7).

ESTIMATING GROWTH PARAMETERS

Continuous rate of increase

We define the continuous rate of increase as the parameter r in the exponential growth SDE (Eq. 12). By adopting this definition, we abandon the notion that there is anything "intrinsic" about a deterministic formulation of exponential growth. Rather, r is simply a constant related to statistical properties of the stochastic process $N(t)$. Specifically, rn represents the infinitesimal mean of $N(t)$, that is, $rn\Delta t$ is approximately the average amount of change in $N(t)$ over a tiny time interval Δt, given that $N(t) = n$. The advantage of this definition becomes apparent when r must be estimated from time series observations.

Any parameter in a deterministic model can only be sensibly estimated from time series data by embedding the model in a statistical/stochastic framework, that is, by converting the model into a stochastic one. The parameter r in the deterministic growth equation $N(t) = n_0 e^{rt}$ is often estimated by linear or nonlinear least squares, with observations on population size serving as the "dependent variable," and observations on time as the "independent variable." The quality of such an estimate, however, depends upon: (a) an assumed statistical model of errors in the regression, and (b) whether that error model adequately represents how the data arise. The usual regression package printouts of confidence intervals for r based on least squares calculations assume independent, normal errors. This amounts to no more than fitting a stochastic growth model to the data, and a bad one at that, since the observations in a population time series are seldom independent.

Instead, the exponential growth SDE (Eq. 12) provides at the outset an explicit, realistic structure of dependence in observations of a population's size through time. Under the exponential growth SDE, log-scale increments of population size are independent, but the population's actual sizes are not. This dependence structure is directly incorporated into parameter estimates through the likelihood function (Eq. 22). If the model is an acceptable representation of the system (as evaluated by the diagnostic procedures described in the *Model evaluation* section, above), then the parameter r as defined here can be efficiently estimated from data.

The quantity r is a function, given by Eq. 13, of the parameters μ and σ^2. Substituting the ML estimates of μ and σ^2 in Eq. 13 produces the ML estimate of r:

$$\hat{r} = \hat{\mu} + (\hat{\sigma}^2/2). \qquad (57)$$

The ML estimate is biased due to the bias of $\hat{\sigma}^2$, though the bias becomes negligible when q is large. An unbiased estimate of r results from using $\tilde{\sigma}^2$ instead of $\hat{\sigma}^2$:

$$\tilde{r} = \hat{\mu} + (\tilde{\sigma}^2/2). \qquad (58)$$

Since $\hat{\mu}$ and $\tilde{\sigma}^2$ are "sufficient statistics" (all information in the data about μ and σ^2 is contained in $\hat{\mu}$ and $\tilde{\sigma}^2$), a fundamental result in statistics known as the Rao-Blackwell theorem applies (see Rice 1988:261): $\tilde{r}$ is the uniformly minimum variance unbiased (UMVU) estimate of r. In other words, no other unbiased estimate of r has a smaller variance. Curiously, $\hat{r}$ has a smaller variance than $\tilde{r}$, due to the well-known fact that $\hat{\sigma}^2$ has a smaller variance than $\tilde{\sigma}^2$. Though $\hat{r}$ will be, on the average, "closer" to r than $\tilde{r}$, it will tend to underestimate r.

The variance of $\tilde{r}$ is just the sum of the variances of $\hat{\mu}$ and $\tilde{\sigma}^2/2$, since $\hat{\mu}$ and $\tilde{\sigma}^2$ are independent:

$$\text{Var}(\tilde{r}) = \text{Var}(\hat{\mu}) + \text{Var}(\tilde{\sigma}^2/2)$$
$$= (\sigma^2/t_q) + \{\sigma^4/[2(q-1)]\}. \qquad (59)$$

The distribution of $\tilde{r}$ is complicated; it is the sum of a normal-distributed random variable, $\hat{\mu}$, and a gamma-distributed (=constant·chi-square) random variable, $\tilde{\sigma}^2/2$. However, the distribution is fairly well approximated by a normal distribution for moderately large values of q (say, 20 or so):

$$\tilde{r} \sim \text{normal}\left(r, \frac{\sigma^2}{t_q} + \frac{\sigma^4}{2(q-1)}\right). \qquad (60)$$

An approximate $100(1 - \alpha)\%$ confidence interval for r is thus given by

$$\left(\tilde{r} \pm z_{\alpha/2} \sqrt{\tilde{\sigma}^2\left[\frac{1}{t_q} + \frac{\tilde{\sigma}^2}{2(q-1)}\right]}\right). \qquad (61)$$

Here $z_{\alpha/2}$ is the $100[1 - (\alpha/2)]$th percentile of the standard normal distribution.

We point out that the symbol r in Braumann's (1983b) paper on estimation corresponds to the exponential growth SDE (Eq. 12) as defined by the Stratonovich stochastic integral. Braumann noted that the ML estimate of r he derived does not depend on q (the number of observations), but just on t_q (the amount of time the system has been observed). In fact, under the Stratonovich interpretation of the SDE (Eq. 12), r equals μ (the infinitesimal mean of the log-scaled process), and the infinitesimal mean of $N(t)$ is $r + (\sigma^2/2)$. In this case, the ML estimate of r is given by Eq. 24 for the ML estimate of μ.

The fundamental discrete-time nature of population growth for many vertebrate species (e.g., seasonal breeding periods), coupled with the limit theorems on stochastic projection matrices (Tuljapurkar and Orzack 1980, Heyde and Cohen 1985), suggests instead that the Ito interpretation of the SDE (Eq. 12) be used in typical endangered species contexts. Estimates of r (Eqs. 57 and 58) then become strongly dependent on q as well as t_q. Braumann (1983a) rightly observed that the differing definitions of the SDE (Eq. 12) produce semantic differences in how r is viewed as a central tendency measure. Under the Stratonovich calculus r is the rate constant in the geometric mean (Eq. 11) of $N(t)$, while under the Ito calculus r is the rate constant in the mean (Eq. 9) of $N(t)$. The mean of a lognormal random variable reflects the skewness of the distribution. The small but real possibility of large values of $N(t)$ has a strong upward influence on the mean, and so the lognormal shape parameter σ^2 appears in Eq. 9 for the mean. The estimates of r presented here (Eqs. 57 and 58) thus depend on the information available for estimating σ^2 (i.e., q) as well as μ (i.e., t_q).

Finite rate of increase

We define the finite rate of increase, denoted λ, as follows:

$$\begin{aligned} \lambda &= \exp(r) = \psi(1; n_0, \mu, \sigma^2)/n_0 \\ &= \exp[\mu + (\sigma^2/2)]. \end{aligned} \qquad (62)$$

It is the mean population size after 1 yr (Eq. 9) divided by the initial size. The quantity approximates the dominant eigenvalue of the average projection matrix of the population, when the stochastic projection matrices are serially uncorrelated (see Eq. 3). More generally, λ is simply a positive constant appearing when the SDE (Eq. 12) is rewritten in discretized form as a stochastic difference equation:

$$N(t + 1) = \lambda N(t)L, \qquad (63)$$

where L is a lognormal random variable with a mean of 1 and a shape parameter σ^2 (i.e., $\log L \sim \text{normal}[-(\sigma^2/2), \sigma^2]$). The distributions of $N(t)$ in Eqs. 12 and 63 coincide for integer values of t. A more descriptive term for λ might be the "discrete rate of increase," but the above term is widely used in the context of deterministic growth models.

The ML estimate of λ is obtained from the definition (Eq. 62) using the ML estimates of μ and σ^2:

$$\hat{\lambda} = \exp(\hat{r}) = \exp[\hat{\mu} + (\hat{\sigma}^2/2)]. \qquad (64)$$

An alternate estimate uses the unbiased $\tilde{r}$ instead of $\hat{r}$: $\exp(\tilde{r})$. Both of these are biased estimates of λ, though the bias disappears as q and t_q become large. The bias can be eliminated with a little programming effort. Shimizu and Iwase (1981; see also Shimizu 1988) studied estimation of functions in the form $\exp(a\mu + b\sigma^2)$, where a and b are constants. With the help of their results, we find that the following expression gives the UMVU estimate of λ:

$$\tilde{\lambda} = \exp(\hat{\mu})_0F_1\left(\frac{q-1}{2}; \frac{q-1}{4}\hat{\sigma}^2\right). \qquad (65)$$

Here $_0F_1(v; z)$ is the "zero-F-one" hypergeometric function, an easily computed infinite series:

$$_0F_1(v; z) = \sum_{j=0}^{\infty} \frac{z^j}{(v)_j(j!)}, \qquad (66)$$

where $(v)_j$ denotes $v(v + 1) \ldots (v + j - 1)$, with $(v)_0 = 1$. Successive terms in the series are handily calculated with a recurrence relation: writing $Q_j = z^j/[(v)_j(j!)]$, we see that $Q_{j+1} = Q_jz/[(v + j)(j + 1)]$ and $Q_0 = 1$. The terms rapidly become small, and the sum can be truncated when adding more terms produces negligible change.

Additionally, Shimizu and Iwase's (1981) results allow us to obtain the variance of $\hat{\lambda}$:

$$\text{Var}(\tilde{\lambda}) = \lambda^2\left[\exp\left(\frac{\sigma^2}{q}\right)_0F_1\left(\frac{q-1}{2}; \frac{(q-1)^2}{4q^2}\sigma^4\right) - 1\right]. \qquad (67)$$

The distribution of $\tilde{\lambda}$ will converge to a normal$[\lambda, \text{Var}(\tilde{\lambda})]$ distribution as q and t_q become large. Estimates of μ and σ^2 can be substituted into the variance formula (Eq. 67) for constructing confidence intervals of the form $\{\tilde{\lambda} \pm z_{\alpha/2}[\widehat{\text{Var}}(\tilde{\lambda})]^{1/2}\}$. However, the distribution of

$\tilde\lambda$ is skewed, and convergence to normality might be slow. The situation should be studied further with computer simulations. In the interim, we recommend instead the use of the approximate normal distribution of $\tilde r$ for constructing approximate $100(1 - \alpha)\%$ confidence intervals for λ:

$$\left(\exp\left\{\tilde r \pm z_{\alpha/2}\sqrt{\tilde\sigma^2\left[\frac{1}{t_q} + \frac{\hat\sigma^2}{2(q-1)}\right]}\right\}\right). \quad (68)$$

Mean population size

The mean population size given by Eq. 9 represents the expected value of $N(t)$ at a time t. This quantity is also a function of the unknown parameters μ and σ^2, and its ML estimate becomes

$$\hat\psi = \psi(t; n_0, \hat\mu, \hat\sigma^2) = n_0\exp(\hat r t) = n_0\hat\lambda^t$$
$$= n_0 \exp\{[\hat\mu + (\hat\sigma^2/2)]t\}. \quad (69)$$

As usual, the ML estimate retains a small-sample bias, as does the estimate defined using $\tilde r$ instead of $\hat r$. But the function is in the form constant$\cdot\exp(a\mu + b\sigma^2)$, and Shimizu and Iwase's (1981) results can be applied directly to obtain the UMVU estimate of the mean population size:

$$\tilde\psi = n_0\exp(\hat\mu t)_0F_1\left(\frac{q-1}{2}; \frac{t(q-t)}{4}\hat\sigma^2\right). \quad (70)$$

The variance of this estimate is found to be

$$\mathrm{Var}(\tilde\psi) = n_0^2\exp(2\mu t + \sigma^2 t)$$
$$\cdot\left[\exp\left(\frac{t^2\sigma^2}{q}\right)_0F_1\left(\frac{q-1}{2}; \frac{t^2(q-t)^2}{4q^2}\sigma^4\right) - 1\right]. \quad (71)$$

We note that Eqs. 70 and 71 reduce to 65 and 67 when $t = 1$ (except for the constant term, n_0). Approximate interval estimates of $\psi(t; n_0, \mu, \sigma^2)$ can be constructed from the fact that the distribution of $\tilde\psi$ converges to a normal$[\psi(t; n_0, \mu, \sigma^2), \mathrm{Var}(\tilde\psi)]$ distribution. But again, the quality of the distribution approximation has not been studied. Instead, the interval given by

$$\left(n_0\exp\left\{\tilde r t \pm z_{\alpha/2}t\sqrt{\tilde\sigma^2\left[\frac{1}{t_q} + \frac{\hat\sigma^2}{2(q-1)}\right]}\right\}\right) \quad (72)$$

should provide an acceptable approximation of a $100(1 - \alpha)\%$ confidence interval for the mean population size.

Geometric mean population size

The geometric mean of $N(t)$ (and the median of the lognormal transition pdf) given by Eq. 11 is a function of just one unknown parameter, μ. Its ML estimate is

$$\hat\beta = \beta(t; n_0, \hat\mu) = n_0\exp(\hat\mu t). \quad (73)$$

The ever-present small-sample bias can be removed as before with Shimizu and Iwase's (1981) results. The UMVU estimate of $\beta(t; n_0, \mu)$ is found to be

$$\tilde\beta = n_0\exp(\hat\mu t)_0F_1\left(\frac{q-1}{2}; \frac{t^2}{4}\hat\sigma^2\right). \quad (74)$$

The variance of $\tilde\beta$ becomes

$$\mathrm{Var}(\tilde\beta) = n_0^2\exp(2\mu t)$$
$$\cdot\left[\exp\left(\frac{t^2\sigma^2}{q}\right)_0F_1\left(\frac{q-1}{2}; \frac{t^4}{4q^4}\sigma^4\right) - 1\right]. \quad (75)$$

While approximate confidence intervals for $\beta(t; n_0, \mu)$ could be constructed from the asymptotic normal$[\beta(t; n_0, \mu), \mathrm{Var}(\tilde\beta)]$ distribution of $\tilde\beta$, using the exact confidence intervals available for μ (Eq. 31) is more straightforward. A $100(1 - \alpha)\%$ confidence interval for $\beta(t; n_0, \mu)$ is

$$\{n_0\exp[(\hat\mu \pm t_{\alpha/2,q-1}\tilde\sigma/\sqrt{t_q})t]\}. \quad (76)$$

Tuljapurkar (1982a) defined a quantity, denoted α, which is the (finite or discrete) growth rate of the geometric mean population size. For the diffusion process $N(t)$, that quantity is

$$\alpha = \beta(1; n_0, \mu)/n_0 = \exp(\mu). \quad (77)$$

Just as the geometric mean $\beta(t; n_0, \mu)$ characterizes "typical" sample paths of $N(t)$ better than does the mean $\psi(t; n_0, \mu, \sigma^2)$, the quantity α gives a better portrait of the growth rate of those sample paths than does λ. Eqs. 73–76 all apply directly to estimation of α, just by setting $t = 1$ and dividing by n_0 (or n_0^2 in Eq. 75). In particular,

$$\hat\alpha = \beta(1; n_0, \hat\mu)/n_0 = \exp(\hat\mu) \quad (78)$$

is the ML estimate, and

$$\tilde\alpha = \exp(\hat\mu)_0F_1\left(\frac{q-1}{2}; \frac{\hat\sigma^2}{4}\right) \quad (79)$$

is the UMVU estimate.

We mention that the UMVU estimates of λ and α we have calculated so far in practice (see *Examples*) differed only slightly from the respective ML estimates. The bias in the ML estimates may turn out to be negligible upon further study.

Forecasting

The essential question in forecasting is as follows: given observations $n_0, n_1, \ldots, n_q$ of the system at times $t_0, t_1, \ldots, t_q$, what is the best prediction for the state of the system at some future time $t > t_q$? We will focus on predicting the value of $X(t) = \log N(t)$, since the mean of $X(t)$ [and the geometric mean of $N(t)$] typifies the system better than the mean of $N(t)$. The Markov property of the diffusion process implies that the expected value of $X(t)$, given that $X(t_q) = x_q$, does not depend on the earlier observations $x_0, x_1, \ldots, x_{q-1}$. Also, the stationary transitions property implies that the expected value depends on the amount of time, $s = t - t_q$, elapsed since the last observation, but not otherwise on t_q. Thus we have

$$E[X(t)|X(t_q) = x_q] = \gamma(s; x_q, \mu) = x_q + \mu s. \quad (80)$$

The ML estimate of $\gamma(s; x_q, \mu)$, found by substituting $\hat{\mu}$, is also the UMVU estimate:

$$\hat{\gamma} = \gamma(s; x_q, \hat{\mu}) = x_q + \hat{\mu} s. \quad (81)$$

This is the basic predicted value of $X(t)$.

The accuracy of this prediction can be measured with the mean squared prediction error, that is, the unconditional expectation of $[X(t) - \hat{\gamma}]^2$ over all possible realizations of the process. The expectation is easily calculated by first conditioning on $x_0, x_1, \ldots, x_q$ (and thus on $\hat{\mu}$), and subsequently averaging with respect to the distribution of $\hat{\mu}$. The result becomes

$$E[(X(t) - \hat{\gamma})^2]$$
$$= E\{E[(X(t) - \hat{\gamma})^2 | X(0) = x_0, \ldots, X(t_q) = x_q]\}$$
$$= E\{E[(X(t) - \hat{\gamma} + (x_q + \mu s) - (x_q + \mu s))^2 | \ldots \text{etc.}]\}$$
$$= E\{\sigma^2 s + (\hat{\mu} - \mu)^2 s^2\} = \sigma^2 s\left(1 + \frac{s}{t_q}\right). \quad (82)$$

Provided the model is adequate, the predictor $\hat{\gamma}$ is the best in the sense of minimizing the mean squared error among all linear unbiased predictors.

Prediction intervals for $X(t)$ are based on the fact that $X(t) - \hat{\gamma}$ has a (unconditional) normal distribution with a mean of zero and a variance given by Eq. 82. The quantity $[X(t) - \hat{\gamma}]/[\hat{\sigma}^2 s(1 + (s/t_q))]^{1/2}$ then has a Student's t distribution with $q - 1$ df. The resulting $100(1 - \alpha)\%$ prediction interval for $X(t)$ is

$$\left[\hat{\gamma} \pm t_{\alpha/2, q-1} \sqrt{\hat{\sigma}^2 s\left(1 + \frac{s}{t_q}\right)}\right]. \quad (83)$$

A $100(1 - \alpha)\%$ prediction interval for $N(t)$ is provided by transforming the endpoints of the interval (Eq. 83) with the exponential function $[\exp(\cdot)]$.

ESTIMATING EXTINCTION PARAMETERS

Probability of extinction

The probability (Eq. 15) of attaining a lower threshold is, like other extinction-related quantities, a function of the two unknown parameters, μ and σ^2. Its ML estimate is

$$\hat{\pi} = \pi(x_d, \hat{\mu}, \hat{\sigma}^2)$$
$$= \begin{cases} 1, \hat{\mu} \leq 0; \\ \exp(-2\hat{\mu}x_d/\hat{\sigma}^2), \hat{\mu} > 0. \end{cases} \quad (84)$$

If the estimated probability is <1 (that is, if $\hat{\mu} > 0$), plotting it as a function of threshold population size is sometimes informative. The most recent population size n_q can be taken as the "initial" size by letting $x_d = \log(n_q/n_e)$ in Eq. 84 (see Eq. 14). Then

$$\hat{\pi} = (n_e/n_q)^{2\hat{\mu}/\hat{\sigma}^2}, \quad (85)$$

and this estimated probability of reaching n_e from n_q can be plotted as a function of n_e (or, perhaps $n_q - n_e$)

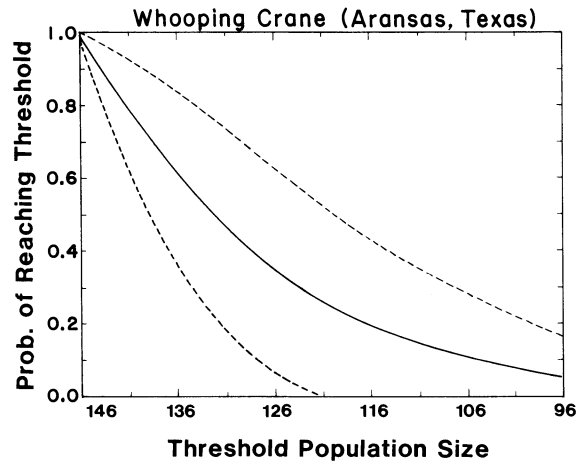**Whooping Crane (Aransas, Texas)**

FIG. 2. Probability of reaching a lower threshold population size plotted as a function of threshold size, as estimated for the Aransas/Wood Buffalo population of Whooping Cranes. Solid line represents maximum likelihood estimate; dashed lines delimit 95% confidence interval.

to study how fast the probability decreases as n_e decreases (Fig. 2).

Expressions for variances of extinction-parameter estimates must at this time be based on approximation results. We relied on the δ method for obtaining such expressions (see Rice 1988:142, Serfling 1980:117). The method essentially involves linearizing the parameter function using a Taylor series expansion. The quality of the approximation varies from function to function, and conducting large-scale simulation studies of these results is an important task for further research. In the meantime, the δ method provides working expressions that must be regarded somewhat cautiously.

Alternative variance estimates could be calculated through bootstrapping or jackknifing (Efron 1982, Lele 1991), or through Monte Carlo simulation; we are studying these possibilities. One method of calculating bootstrap confidence intervals involves using the estimated distributions of $\hat{\mu}$ and $\hat{\sigma}^2$. From Eqs. 29 and 30, the estimated distribution of $\hat{\mu}$ is a normal($\hat{\mu}$, $\hat{\sigma}^2/t_q$) distribution, and the estimated distribution of $\hat{\sigma}^2$ is that of a chi-square($q - 1$) random variable multiplied by $\hat{\sigma}^2/q$. One generates a pair, $\hat{\mu}_B$, and $\hat{\sigma}^2_B$, from these distributions and then calculates the parameter of interest, say $\hat{\pi}_B$, with the pair only if $\hat{\mu}_B$ and $\hat{\sigma}^2_B$ both fall within their respective $100(1 - \alpha)\%$ confidence intervals (Eqs. 31 and 32). One repeats the process hundreds of times or more; the resulting set of values of $\hat{\pi}_B$ range across a $100(1 - \alpha)\%$ confidence interval for $\pi(x_d, \mu, \sigma^2)$. This bootstrapping method is illustrated below in connection with the distribution of extinction time. We mention that the proper observations for deletion in jackknifing are transitions, according to the theory of estimating equations (Lele 1991; discussed in the context of stochastic population models by Dennis 1989b).

Also worth mentioning is the approach of computing profile likelihoods and associated interval estimates (e.g., Kalbfleisch 1986), which bypasses the problem of variance estimation altogether.

We find, in general, the extinction parameters discussed in this section to be more poorly estimated than the growth parameters of the previous section, in that estimated variances calculated for practical examples tend to be extremely large (see *Examples* section). There are various possible explanations for this. First, time series data may intrinsically provide less information for estimating extinction parameters than for growth parameters. Second, when $\mu = 0$, the Wiener-drift process possesses so-called null recurrent behavior; although the probability of ultimately attaining a threshold is 1, the mean time to attain the threshold is infinite (see Eq. 18). Examples we have analyzed typically have estimates of μ near 0, with consequent poor performance of estimates of quantities related to the inverse Gaussian distribution. Finally, the asymptotic properties of the δ-method variance estimates may be poor, though our preliminary results with bootstrapping and jackknifing (not reported here) have yielded comparably sized variance estimates.

The approximate variance of $\hat{\pi}$, when the true value of μ is positive, is found to be

$$\text{Var}(\hat{\pi}) \approx [2x_d\pi(x_d, \mu, \sigma^2)]^2$$
$$\cdot \left\{ \frac{1}{\sigma^2 t_q} + \frac{2\mu^2(q-1)}{(\sigma^2 q)^2} \right\}. \qquad (86)$$

The distribution of $\hat{\pi}$ will converge to a normal distribution with a mean of $\pi(x_d, \mu, \sigma^2)$ and a variance given by Eq. 86, provided $\mu > 0$. However, an asymptotic confidence interval for $\pi(x_d, \mu, \sigma^2)$ based on transforming a confidence interval for the quantity $2\mu x_d/\sigma^2$ in the exponent of Eq. 84 is probably better. Accordingly, an approximate $100(1 - \alpha)\%$ confidence interval for $\pi(x_d, \mu, \sigma^2)$, assuming $\mu > 0$, would take the form

$$\exp\left(-(2\hat{\mu}x_d/\hat{\sigma}^2) \pm z_{\alpha/2}\sqrt{\widehat{\text{Var}}(2\hat{\mu}x_d/\hat{\sigma}^2)} \right). \qquad (87)$$

Here $\text{Var}(2\hat{\mu}x_d/\hat{\sigma}^2) = (4x_d^2/\sigma^2)[(1/t_q) + (2(q-1)\mu^2/(q^2\sigma^2))]$ is the approximate variance of $2\hat{\mu}x_d/\hat{\sigma}^2$, and $\widehat{\text{Var}}$ denotes evaluating the variance at the ML estimates of μ and σ^2. The upper bound of the confidence interval can be taken as 1 when the computed Expression (87) exceeds this value.

Distribution of extinction time

Given that the threshold will be attained, the probability of attaining it before a fixed time t is given by the cdf (Eq. 16) of the inverse Gaussian distribution. This probability is a function of μ and σ^2. Its ML estimate becomes

$$\hat{G} = G(t; x_d, \hat{\mu}, \hat{\sigma}^2). \qquad (88)$$

We point out here that the Expression (16) for the inverse Gaussian cdf in terms of the standard normal cdf, while mathematically correct, can be numerically difficult to evaluate. The problem is numerical overflow of the exponential function [$\exp(\cdot)$] in the second term of the expression. We provide in the *Appendix* an easily programmed algorithm for calculating the inverse Gaussian cdf. Evaluating the cdf at $\hat{\mu}$ and $\hat{\sigma}^2$ then yields the ML estimate (Eq. 88).

Since $G(t; x_d, \mu, \sigma^2)$ is always between 0 and 1, we recommend first constructing a confidence interval for the (logit transform) quantity $H(t; x_d, \mu, \sigma^2) = \log\{G(t; x_d, \mu, \sigma^2)/[1 - G(t; x_d, \mu, \sigma^2)]\}$ and then back-transforming to obtain a confidence interval for $G(t; x_d, \mu, \sigma^2)$. According to the δ method, the large-sample variance of $H(t; x_d, \hat{\mu}, \hat{\sigma}^2)$ is

$$\text{Var}(H(t; x_d, \hat{\mu}, \hat{\sigma}^2))$$
$$\approx \text{Var}(\hat{\mu})(\partial H/\partial\mu)^2 + \text{Var}(\hat{\sigma}^2)(\partial H/\partial\sigma^2)^2$$
$$= (\sigma^2/t_q)\left(\frac{\partial G/\partial\mu}{G(1-G)} \right)^2$$
$$+ 2(q-1)(\sigma^2/q)^2\left(\frac{\partial G/\partial\sigma^2}{G(1-G)} \right)^2. \qquad (89)$$

This quantity can be estimated by substituting the ML estimates $\hat{\mu}$ and $\hat{\sigma}^2$. The derivatives in Eq. 89 can be evaluated numerically, e.g., $\partial G/\partial\mu \approx [G(t; x_d, \mu + \epsilon, \sigma^2) - G(t; x_d, \mu, \sigma^2)]/\epsilon$, etc., for some small number ϵ. With $\widehat{\text{Var}}(H(t; x_d, \hat{\mu}, \hat{\sigma}^2))$ denoting the estimated variance, an approximate $100(1 - \alpha)\%$ confidence interval for $G(t; x_d, \mu, \sigma^2)$ becomes

$$\left[1 + \exp\left(-H(t; x_d, \hat{\mu}, \hat{\sigma}^2) \pm z_{\alpha/2}\sqrt{\widehat{\text{Var}}(H(t; x_d, \hat{\mu}, \hat{\sigma}^2))} \right) \right]^{-1}. \qquad (90)$$

As an alternative, obtaining bootstrap confidence intervals using the method discussed earlier (*Estimating extinction parameters: Probability of extinction*) is straightforward. One can calculate the whole function $G(t; x_d, \hat{\mu}_B, \hat{\sigma}^2_B)$, that is, calculate G for a whole range of values of t, for each bootstrap pair $\hat{\mu}_B, \hat{\sigma}^2_B$. The resulting hundreds of functions of t, when plotted on one graph, will shade in a confidence region for $G(t; x_d, \mu, \sigma^2)$.

Plotting the estimated cdf (Eq. 88) and associated confidence intervals/regions as functions of t reveals how the cdf estimate becomes extremely uncertain as t becomes large, if the number of observations is small (Fig. 3). Though the confidence region as estimated for the California Condor is enormous, Fig. 3 nonetheless contains useful information. For instance, it indicates that, in 1980, the probability of the population declining to one bird within 5 yr could have been higher than 0.15.

Mean time to extinction

The mean time (Eq. 18) to reach the threshold, given the threshold is reached, is a function of just one unknown parameter, μ. Its ML estimate is

$$\theta = \theta(x_d, \hat{\mu}) = x_d/|\hat{\mu}|. \qquad (91)$$

Interestingly, when μ is positive, so that attaining the lower threshold is not certain, the mean time to reach the threshold becomes smaller as μ becomes larger (Lande and Orzack 1988). This is because any sample paths that cross the lower threshold usually do so quickly when the growth rate of the process is high. A small estimated mean time to extinction can thus occur in a rapidly growing population as well as a rapidly declining one. The difference between these two situations will be reflected in the estimated probability (Eq. 84) of reaching the threshold. While a negative estimate of μ predicts certain attainment of the threshold, a large (relative to $\hat{\sigma}^2$) positive estimate of μ predicts that attaining the threshold at all is an unlikely event.

The mean and the variance of $\hat{\theta}$ are infinite (i.e., do not exist). However, the distribution of $\hat{\theta}$ converges to a normal distribution, which possesses a mean and a variance. The mean of the asymptotic normal distribution is $\theta(x_d, \mu)$, and the variance is given by the δ method:

$$\text{Var}(\hat{\theta}) \approx x_d^2\sigma^2/(\mu^4 t_q). \qquad (92)$$

Thus, probability statements about the estimated mean time to extinction can be approximated with calculations involving a normal distribution, even though the estimate itself has a distribution without moments. In particular, an approximate $100(1 - \alpha)\%$ confidence interval for $\theta(x_d, \mu)$ would be

$$\left(\hat{\theta} \pm z_{\alpha/2}\sqrt{\widehat{\text{Var}(\hat{\theta})}}\right). \qquad (93)$$

Median time to extinction

The median of the highly skewed inverse Gaussian distribution is probably a more representative measure of central tendency. The mean is inflated by the rare sample paths of $N(t)$ that take enormous amounts of time to reach the threshold. The median, denoted $\xi_{0.5}(x_d, \mu, \sigma^2)$, is defined as a root of a nonlinear equation (Eq. 20); the implicit function theorem (e.g., Rudin 1964:195) guarantees that the median is locally a differentiable function of μ and σ^2. The ML estimate of the median would be found as a root of

$$G(\hat{\xi}_{0.5}; x_d, \hat{\mu}, \hat{\sigma}^2) - 0.5 = 0. \qquad (94)$$

This equation must be solved numerically for $\hat{\xi}_{0.5}$ by using an iterative procedure such as Newton's method (see Press et al. 1986:240), coupled with a subroutine for evaluating the inverse Gaussian cdf (*Appendix*). A simple plot of $G(t; x_d, \hat{\mu}, \hat{\sigma}^2)$ over a range of values of t affords easy selection of a "close" starting value for the iterations.

The asymptotic variance of $\hat{\xi}_{0.5}$ requires some additional but straightforward computing. The δ method yields
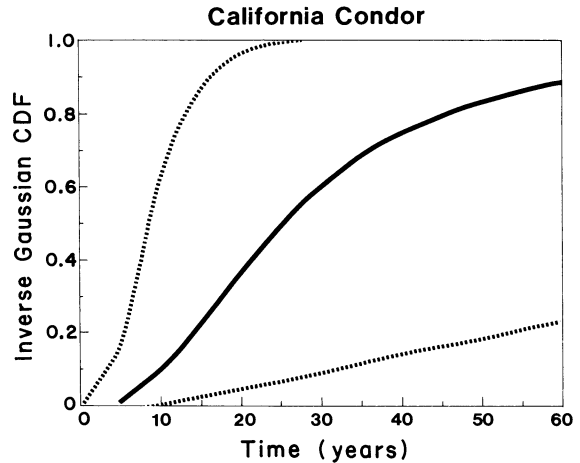
California Condor

FIG. 3. Cumulative distribution function (CDF) of the inverse Gaussian distribution, plotted using maximum likelihood estimates of μ and σ^2 for the California Condor. The distribution is that of the time required for 12 birds to decline to 1 bird. Solid line represents maximum likelihood estimate; dashed lines delimit 95% confidence region calculated with a bootstrap method.

$$\text{Var}(\hat{\xi}_{0.5}) \approx (\sigma^2/t_q)\left(\frac{\partial G/\partial\mu}{g(\xi_{0.5}; x_d, \mu, \sigma^2)}\right)^2$$

$$+ 2(q - 1)(\sigma^2/q)^2\left(\frac{\partial G/\partial\sigma^2}{g(\xi_{0.5}; x_d, \mu, \sigma^2)}\right)^2. \qquad (95)$$

Here $g(\cdot; \cdot, \cdot, \cdot)$ is the inverse Gaussian pdf (Eq. 17); the derivatives can be calculated numerically. The approximate $100(1 - \alpha)\%$ confidence interval based on the asymptotic normal distribution of $\hat{\xi}_{0.5}$ is constructed in the usual manner:

$$\left[\hat{\xi}_{0.5} \pm z_{\alpha/2}\sqrt{\widehat{\text{Var}(\hat{\xi}_{0.5})}}\right]. \qquad (96)$$

All quantities in $\widehat{\text{Var}(\hat{\xi}_{0.5})}$ should be evaluated at $\hat{\xi}_{0.5}$, $\hat{\mu}$, and $\hat{\sigma}^2$.

ML estimates for other percentiles of the inverse Gaussian distribution are calculated as above, using the equation

$$G(\hat{\xi}_p; x_d, \hat{\mu}, \hat{\sigma}^2) - p = 0 \qquad (97)$$

to obtain the $100p$th percentile.

Most likely extinction time

Estimating the location of the mode of the inverse Gaussian distribution is also of interest, since the mode represents the most likely time that the threshold will be attained. From Eq. 21, the ML estimate of the mode is

$$\hat{t}^* = t^*(x_d, \hat{\mu}, \hat{\sigma}^2)$$

$$= \frac{x_d}{|\hat{\mu}|}\left\{\left[1 + \frac{9}{4\hat{\nu}^2}\right]^{1/2} - \frac{3}{2\hat{\nu}}\right\}, \qquad (98)$$

where $\hat{\nu} = x_d|\hat{\mu}|/\hat{\sigma}^2$. The estimate corresponds to the inflection point of the estimated inverse Gaussian cdf

(Eq. 16). The variance of the asymptotic normal distribution of $\hat{t}^*$, given by the δ method, is

$$\text{Var}(\hat{t}^*) \approx (\sigma^2/t_q)(\partial t^*/\partial \mu)^2$$
$$+ 2(q - 1)(\sigma^2/q)^2(\partial t^*/\partial \sigma^2)^2. \quad (99)$$

When estimating this variance, the derivatives can be calculated numerically; or, an enterprising investigator might prefer to obtain analytical expressions. In either case, an approximate $100(1 - \alpha)\%$ confidence interval for $t^*(x_d, \mu, \sigma^2)$ becomes

$$\left[\hat{t}^* \pm z_{\alpha/2}\sqrt{\widehat{\text{Var}(\hat{t}^*)}}\right]. \quad (100)$$

EXAMPLES

We selected seven species for detailed analyses, primarily because long-term population estimates were available for each species. Though six of the species are birds, a variety of ecological conditions and life history strategies are represented. The situations range from a species now extirpated from the wild (the California Condor) to a once extremely threatened species now undergoing a promising recovery (the Whooping Crane). In this section, we briefly review some relevant biological aspects and management efforts for each species, and we discuss estimates and predictions resulting from fitting the model.

Whooping Crane

The Whooping Crane has been the subject of protection efforts by the National Audubon Society and the Federal governments of Canada and the United States of America for more than half a century. The Whooping Crane is a long-lived bird that stands 1.5 m tall with a wingspan of 2.1 m. It becomes sexually mature on average at 5 yr of age and normally lays two eggs per clutch. Usually only one chick is raised to fledging age. One viable wild population exists, which breeds in Wood Buffalo National Park in northwestern Canada and winters at the Aransas National Wildlife Refuge on the gulf coast of Texas. As a result of major research and management programs, this population increased from 18 birds in 1938 to 146 birds in 1989. Efforts to recover the Aransas/Wood Buffalo flock include legal protection, manipulation of hunting seasons for Snow Geese and Sandhill Cranes, egg manipulation, habitat protection, and habitat improvement (USFWS 1986). Predator control programs have also been undertaken at the Grays Lake, Idaho summering area (USFWS 1986). Attempts to create a second wild flock at Grays Lake by placing Whooping Crane eggs in Sandhill Crane nests have been unsuccessful, because no breeding has occurred, and flock mortality is unusually high in comparison to the Aransas/Wood Buffalo population.

An annual census of wintering Whooping Cranes commenced in 1938 and has continued to the present (Fig. 4). The data we analyzed consist of combined annual counts of young birds and birds with adult plumage, as presented by Boyce (1987) and supplemented by recent counts (J. Lewis, *personal communication*). An overall trend of exponential growth, with fluctuations, is evident in the data.

Model parameter estimates indicate a favorable outlook for this population (Tables 1 and 2). We present extinction parameter estimates corresponding to threshold populations of $n_e = 100$ and $n_e = 10$. Thresholds for this and other populations were selected with the idea that management efforts might be planned usefully around the risks of dropping 1 and 2 orders of magnitude (e.g., from 3-digit abundances to 2, from 3 to 1, etc.). The estimates presented here were calculated with the 1940–1941 transition omitted from the data, since the decrease of 26 to 16 birds is a significant outlier ($I_j = -3.75$, $c_{0.05} = 3.2$, $E_j = -4.4$, $j = 3$) according to the outlier procedures (Eqs. 35 and 37). That transition has an inordinate influence on the parameter estimates if included in the data; the UMVU estimate of λ, for instance, decreases from 1.061 to 1.052.

Model diagnostic procedures, once the outlier transition is removed, reveal no significant additional outliers, influential observations, or first-order autocorrelation among residuals (the outlier test [Eq. 35] and the Durbin-Watson test [Eq. 34] were conducted here, and in all examples to follow, at a significance level of $\alpha = .05$). However, a spectral analysis of the residuals confirms a 10-yr cycle in the data reported by Boyce (1987), Boyce and Miller (1985), and Nedelman et al. (1987). Further refinement of the model predictions might be possible by using a time-dependent infinitesimal mean, $\mu(t)$, in the Wiener-drift process (perhaps a sine wave); we are currently developing such an approach.

We also find no evidence using the slope-change test (Eq. 51) that the value of μ changed starting with the 1957–1958 transition ($T_{48} = -0.1804$, $P = .86$). Binkley and Miller (1988) fitted a model in the form log $N(t) = \log N(0) + \mu t + \epsilon$ using ordinary least squares regression, corrected for an error structure having first-order serial correlation. They identified 1957 as a year in which the population started to recover more rapidly, because their analysis indicated a shift in μ occurred. Their model is similar to ours in that $E[\log N(t)]$ is a linear function of time in both models. The "stochastic process" aspects of our model are more explicit, though, allowing for a connection to the stochastic theory of age-structured populations and for estimation of extinction-related quantities.

The Whooping Crane recovery plan (USFWS 1986) calls for downlisting the Whooping Crane from endangered to threatened status when (among other things) 40 nesting pairs are attained in this population. Binkley and Miller (1988) estimate, using a survivorship analysis, that 40 nesting pairs would correspond to a total population of 153 birds.
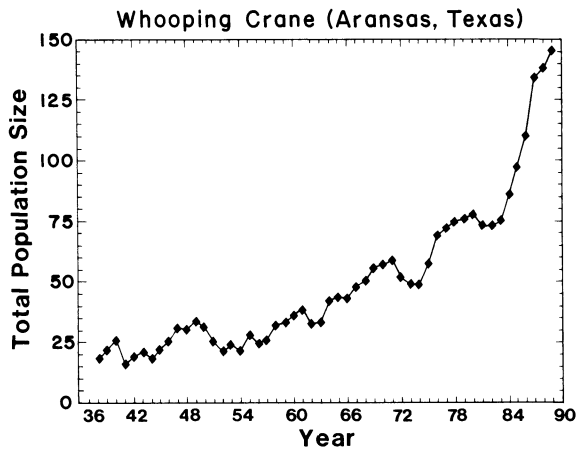
FIG. 4. Total size of the Aransas/Wood Buffalo Whooping Crane population, from 1938–1988. Data are from Boyce (1987), supplemented by more recent counts.

In order to compare Whooping Crane recovery forecasts from Binkley and Miller's (1988) model and our own, we refitted our model to the Whooping Crane data using observations only through 1986, that is, only those observations available to Binkley and Miller. The ML estimates of 4.884×10^{-2} and 1.491×10^{-2} for μ and σ^2, respectively, resulted after deleting the 1940–

1941 transition. According to the ML estimate of the mean time to reach 153 birds from 110 birds (Table 2), the population is expected under our model to attain the recovery threshold by 1993. The prediction is more optimistic than Binkley and Miller's (1988) prediction of 1997. Their estimate arises from projecting the estimated mean population size (in their model) until it hits 153. We note that in stochastic processes, the time until the mean population size hits a threshold is not generally equal to the mean time until the population hits the threshold. Also, the estimated inverse Gaussian distribution here is somewhat skewed; the ML estimates of the 5th, 50th (median), and 95th percentiles are, respectively, 1.31, 4.68, and 19.3 yr (starting from 1986). We note that the population has been progressing rapidly toward the threshold of 153 birds since 1986 (Fig. 4).

Grizzly bear

The grizzly bear was formerly found throughout most of western North America. Within the contiguous 48 states of the United States, it presently occurs in six populations (USFWS 1982, Allendorf and Servheen 1986). The numbers of bears in these populations range from fewer than 10 apiece in the North Cascade and Selway-Bitterroot populations, to 440–680 in the northern continental divide population. A 12-yr mark-

TABLE 1. Estimated growth parameters for the Whooping Crane (WC), grizzly bear (GB), Kirtland's Warbler (KW), California Condor (CC), Puerto Rican Parrot (PP), Palila (PA), and Laysan Finch (LF). Numbers in parentheses are 95% confidence limits.

Animal	$\hat{\mu}$*	$\tilde{\sigma}^2$†	$\tilde{r}$‡	$\tilde{\lambda}$§	$\tilde{\alpha}$‖	q¶	t_q (yr)#
WC**	5.157×10^{-2}	1.475×10^{-2}	5.895×10^{-2}	1.061	1.053	50	50
	(1.705×10^{-2})	(1.029×10^{-2})	(2.516×10^{-2})	(1.025)	(1.017)		
	(8.610×10^{-2})	(2.291×10^{-2})	(9.274×10^{-2})	(1.097)	(1.090)		
GB††	-7.493×10^{-3}	8.919×10^{-3}	-3.034×10^{-3}	0.9968	0.9927	27	27
	(-4.486×10^{-2})	(5.531×10^{-3})	(-3.874×10^{-2})	(0.9620)	(0.9561)		
	(2.987×10^{-2})	(1.675×10^{-2})	(3.267×10^{-2})	(1.033)	(1.030)		
KW	-1.873×10^{-2}	1.673×10^{-2}	-1.037×10^{-2}	0.9893	0.9819	20	38
	(-6.264×10^{-2})	(9.674×10^{-3})	(5.183×10^{-2})	(0.9495)	(0.9393)		
	(2.518×10^{-2})	(3.568×10^{-2})	(3.109×10^{-2})	(1.032)	(1.025)		
CC	-7.685×10^{-2}	0.1199	-1.688×10^{-2}	0.9792	0.9297	15	15
	(-0.2686)	(6.429×10^{-2})	(-0.1977)	(0.8206)	(0.7644)		
	(0.1150)	(0.2983)	(0.1639)	(1.1781)	(1.1218)		
PP‡‡	3.400×10^{-2}	1.341×10^{-2}	4.071×10^{-2}	1.041	1.035	20	20
	(-2.020×10^{-2})	(7.757×10^{-3})	(-1.023×10^{-2})	(0.9898)	(0.9800)		
	(8.820×10^{-2})	(2.861×10^{-2})	(9.165×10^{-2})	(1.096)	(1.092)		
PA	7.732×10^{-2}	0.2190	0.1868	1.190	1.094	9	13
	(-0.2220)	(9.991×10^{-2})	(-8.928×10^{-2})	(0.9146)	(0.8009)		
	(0.3766)	(0.8036)	(0.4629)	(1.589)	(1.457)		
LF	-1.058×10^{-3}	0.3662	0.1820	1.189	1.007	24	22.38
	(-0.2657)	(0.2212)	(-9.010×10^{-2})	(0.9138)	(0.7666)		
	(0.2636)	(0.7206)	(0.4542)	(1.575)	(1.302)		

  * $\hat{\mu}$ = ML (maximum likelihood) estimate of diffusion process drift parameter μ (Eq. 24).
  † $\tilde{\sigma}^2$ = unbiased estimate of diffusion process variance parameter σ^2 (Eq. 26).
  ‡ $\tilde{r}$ = UMVU (uniformly minimum variance unbiased) estimate of continuous rate of increase r (Eq. 58).
  § $\tilde{\lambda}$ = UMVU estimate of finite rate of increase λ (Eq. 65).
  ‖ $\tilde{\alpha}$ = UMVU estimate of geometric finite rate of increase α (Eq. 79).
  ¶ q = number of transitions in data set.
  # t_q = length of time population has been observed.
  ** 1940–1941 transition deleted.
  †† 1983–1984 transition deleted.
  ‡‡ 1972–1972 transition deleted.

TABLE 2.   Estimated extinction parameters for the Whooping Crane (WC), grizzly bear (GB), Kirtland's Warbler (KW), California Condor (CC), Puerto Rican Parrot (PP), Palila (PA), and Laysan Finch (LF), calculated using (maximum likelihood) growth parameter estimates from Table 1. Numbers in parentheses are 95% confidence limits.

Animal	n_q^*	$n_e(n_u)\dagger$	$\hat{\pi}\ddagger$	$\hat{\theta}\S$ (yr)	$\hat{\xi}_{0.5}\|$ (yr)	$\hat{t}^*\P$ (yr)
WC	146	100	6.72×10^{-2}	7.34	5.41	2.82
			(0.00)	(2.60)	(2.76)	(1.88)
			(0.204)	(12.08)	(8.05)	(3.75)
		10	4.90×10^{-9}	52.0	49.4	44.5
			(0.00)	(18.4)	(19.0)	(20.0)
			(7.56×10^{-8})	(85.6)	(79.8)	(68.9)
	110#	(153)	1.00	6.76	4.68	2.18
			$\cdots$	(1.93)	(2.27)	(1.41)
			$\cdots$	(11.6)	(7.08)	(2.95)
GB	47	10	1.00	207	152	79.3
			$\cdots$	(0.00)	(0.00)	(0.00)
			$\cdots$	(1170)	(679)	(179)
		1	1.00	514	448	333
			$\cdots$	(0.00)	(0.00)	(0.00)
			$\cdots$	(2910)	(2280)	(1260)
KW	212	100	1.00	40.1	26.0	11.0
			$\cdots$	(0.00)	(0.00)	(4.39)
			$\cdots$	(126)	(63.1)	(17.5)
		10	1.00	163	143	109
			$\cdots$	(0.00)	(0.00)	(0.00)
			$\cdots$	(512)	(414)	(254)
CC	12	10	1.00	2.37	0.506	9.88×10^{-2}
			$\cdots$	(0.00)	(0.138)	(3.07×10^{-2})
			$\cdots$	(7.60)	(0.875)	(0.167)
		1	1.00	32.3	25.2	14.6
			$\cdots$	(0.00)	(0.00)	(1.80)
			$\cdots$	(104)	(68.8)	(27.4)
PP	38	10	8.04×10^{-4}	39.3	34.5	26.1
			(0.00)	(0.00)	(0.00)	(2.07)
			(9.83×10^{-3})	(96.4)	(78.8)	(50.1)
		1	3.70×10^{-9}	107	102	91.7
			(0.00)	(0.00)	(0.00)	(0.00)
			(1.17×10^{-7})	(263)	(243)	(205)
	23	10	1.17×10^{-2}	24.5	20.1	13.0
			(0.00)	(0.00)	(0.00)	(3.58)
			(9.39×10^{-2})	(60.1)	(44.2)	(22.5)
PA	4358	1000	0.311	19.0	10.5	3.58
			(0.00)	(0.00)	(0.00)	(0.00)
			(1.00)	(78.1)	(29.3)	(6.58)
		100	4.98×10^{-2}	48.8	36.9	20.2
			(0.00)	(0.00)	(0.00)	(0.00)
			(0.531)	(200)	(124)	(42.4)
LF	9349	1000	1.00	2110	30.8	4.75
			$\cdots$	(0.00)	(0.00)	(2.12)
			$\cdots$	(4.92×10^5)	(143)	(7.37)
		100	1.00	4290	125	19.6
			$\cdots$	(0.000)	(0.000)	(8.73)
			$\cdots$	(9.99×10^5)	(1020)	(30.4)

* n_q = starting population size.
† n_e = threshold population size.
‡ $\hat{\pi}$ = ML (maximum likelihood) estimate of probability of attaining threshold (Eq. 84).
§ $\hat{\theta}$ = ML estimate of (conditional) mean time to reach threshold (Eq. 91).
‖ $\hat{\xi}_{0.5}$ = ML estimate of (conditional) median time to reach threshold.
¶ $\hat{t}^*$ = ML estimate of (conditional) most likely time to reach threshold.
# No observations used after 1986.

recapture study of the Yellowstone population gave a peak population size estimate of 245 bears in 1967 (Craighead et al. 1974). The isolated Yellowstone population is the subject of our analysis.

Since the Craighead et al. (1974) study, reliable, periodic estimates of total population size have not been available. However, Knight and Eberhardt (1984, 1985) and Eberhardt et al. (1986) devised an estimate of the minimal number of fully adult females in the population during the 1st yr of any 3-yr period. The estimate

is simply a running 3-yr sum of the numbers of female grizzlies observed with cubs. The estimate is based on two points: females with cubs are probably the most observable segment of the population in aerial surveys, and the breeding interval is at least 3 yr. The resulting time series of these estimates, calculated from data published by Eberhardt et al. (1986) and from more recent unpublished data (R. R. Knight, *personal communication*), shows substantial fluctuations (Fig. 5). We mention that the Tuljapurkar and Orzack (1980) results, giving an approximate normal transition distribution to log $N(t)$, apply in general to any linear combinations of age or stage classes. Thus, segments of a population can be analyzed with the model.

Model parameter estimates for this set of data suggest that the Yellowstone grizzly population is doomed to extinction, though not in our lifetimes (Tables 1 and 2). The confidence interval for μ contains positive values (Table 1), indicating that such an interval estimate of the extinction probability would contain values <1. However, the high value of σ^2 creates a large estimated chance of extinction even if μ is slightly positive. The estimated distribution of the time required to reach a threshold of 10 bears, portrayed earlier in Fig. 1, attaches nontrivial likelihood to times ranging over hundreds of years. A level of 10 adult female bears would represent, if not extinction, a serious failure of management and preservation efforts.

The estimates given in Table 1 were calculated with the 1983–1984 transition deleted. In 1986, the observed number of females with cubs in the population jumped from 9 to 25 (R. R. Knight, *personal communication*). This large increase caused an anomalous increase in the 1983–1984 transition (the 3-yr sum jumps from 39 to 51). The residual screening procedures tag the transition as a suspect outlier; specifically, for that transition $E_j = 2.9$ (see Eq. 37), a value considerably larger than the values for remaining transitions. If the transition is included in the analysis, the estimate of μ becomes slightly positive ($\hat{\mu} = 0.002356$), the estimate of σ^2 becomes larger ($\hat{\sigma}^2 = 0.01130$), and the estimated probability of reaching 10 bears, though smaller, is still high enough to cause concern ($\hat{\pi} = 0.51$). Preservation strategies that bank upon such fortuitous increases in population size are not likely to inspire confidence.

Diagnostic indicators are otherwise acceptable, once the 1983–1984 transition is omitted. No significant outliers are present, according to the outlier test (Eq. 35), and the influence measures, Eqs. 38, 39, suggest there are no highly influential transitions. The test for first-order autocorrelation (Eq. 34) is inconclusive.

Closure of garbage dumps in Yellowstone National Park during 1970–1971 was hypothesized to have substantial negative effects on the grizzly population. Such effects have been reported for mortality and other demographic parameters (Knight and Eberhardt 1984). We find no evidence of a change in μ starting with the
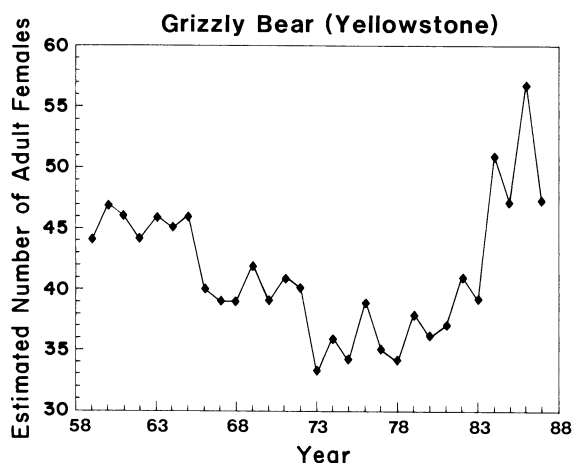

FIG. 5. Estimated number of adult females in the Yellowstone National Park grizzly bear population, 1959–1987. Data, listed by Eberhardt et al. (1986) and supplemented by recent figures, consist of a 3-yr moving sum of the yearly number of adult females seen with cubs.

1972–1973 transition ($T_{25} = 0.00838$, $P = .99$), using the slope-change test (Eq. 51). However, we do find evidence of a change in σ^2 ($F_{11,13} = 0.158$, $P < .01$), using the variance-change test (Eq. 56). In fitting the two models, the 1965–1966 transition was tagged as an outlier ($|E_j| > 2.9$) and deleted from the analysis in addition to the 1983–1984 transition. Results thus indicate greater variability in the female population size index commencing around the time of the dump closures. Two possible explanations for this variability increase are: (a) without access to the stable food supply provided by the dumps, the bears may be more influenced by fluctuations in important wild food sources such as whitebark pine nuts and ungulate carcasses; (b) the method of data collection changed from observations at dumps (Craighead et al. 1974) to aerial observations (Knight and Eberhardt 1985), which may have resulted in decreased precision. It will be interesting to determine, after more years of data accumulate, if any post-1988 fire effects can be detected with the model.

Several stochastic, age-structured simulation models of the Yellowstone grizzly population have been constructed (Shaffer 1978, Knight and Eberhardt 1984, 1985, Shaffer and Samson 1985, Suchy et al. 1985, Eberhardt et al. 1986). As best we can ascertain, these models conform to the Tuljapurkar and Orzack (1980) criteria, and the log-transformed total population size projected by these models should have the approximate statistical characteristics of a Wiener-drift process. A standing research problem is to study, using extensive simulations, how well the Wiener-drift process (and associated statistical inferences) portrays the output of these detailed projection models.

Kirtland's Warbler

The population of the Kirtland's Warbler has fluctuated around 200 singing males since 1971, after a
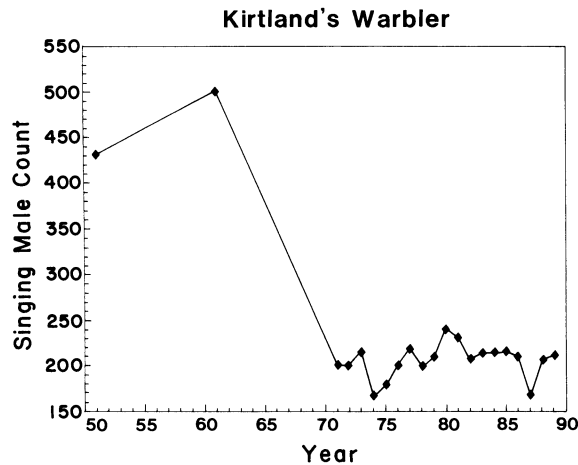
Kirtland's Warbler

FIG. 6. Total count of Kirtland's Warbler singing males, 1951–1989. Data are from Walkinshaw (1983), supplemented by more recent counts.

precipitous population decline occurred sometime during the 1960s (Fig. 6; data from Walkinshaw 1983, and C. B. Kepler, *personal communication*). The decline triggered yearly monitoring counts of singing males in the species' breeding grounds in Michigan, which have continued to the present. Major recovery efforts on behalf of this species have included legal protection of the bird and its habitat, habitat improvement, and control of nest parasites. Limiting factors other than cowbird parasitism have not been well documented (USFWS 1985).

The model parameter estimates calculated for the singing male data give little cause for optimism (Tables 1 and 2). The results indicate that the population could persist above 10 singing males for some hundreds of years, but that the fluctuations will ultimately exterminate the population. In particular, the precipitous drop of the 1960s is well within the bounds of usual variability given by the model, since the drop occurred over a relatively extended period of 10 yr. No significant outliers, influential transitions, or first-order autocorrelation is apparent from analysis of residuals.

California Condor

The California Condor had been regarded as one of the most endangered bird species in North America (Koford 1953, Wilbur 1980, Ogden 1985). By 1987, all the condors remaining in the wild had been taken into captivity, and a captive flock of 32 birds now exists, with 11 potential breeding pairs. It is the largest North American bird, with a wingspan of 2.7 m. It lays a single egg per clutch and is not sexually mature until 5–7 yr of age. The efforts since 1980 to save the wild condor population included nest site protection, supplemental feeding, habitat protection, and a public education program to prevent loss through shooting (USFWS 1984*a*). In 1986, lead poisoning from ingestion of bullets in hunter-killed carcasses was identified

as a major source of mortality (Wiemeyer et al. 1988). Future release programs, in order to have a reasonable chance of success, will have to be limited to lead-free areas where supplemental feeding can be used to minimize exposure to lead-contaminated carcasses. Present recovery efforts focus on increasing the size of the captive flock prior to any release, and on development of methods for releasing birds back into the wild.

Data collected from 1965 onward indicate an inexorable population decline throughout the 1970s (Fig. 7). These data, arising from the dubious "October surveys," are problematic (Wilbur 1980). Astonishingly, no really accurate count was undertaken until the 1980s (Snyder and Johnson 1985). Nonetheless, we take a retrospective look at the October survey data, in order to discover what conclusions could have been drawn with the model in 1980 on the basis of the admittedly poor information available. We follow Snyder and Johnson (1985) in using the maximum number in each multiday October survey as reported by Wilbur (1980) for the population time series.

According to the fitted model, extinction was imminent (Tables 1 and 2). The estimated mean time of 32 yr to decline from 12 birds to 1 bird might by itself have indicated that managers had enough time to attempt to reverse the downward trend of the species in the field. However, the estimated inverse Gaussian cdf gives the probability of extinction within 20 yr to be almost 0.4 (Fig. 3), and the most likely time of extinction to be under 15 yr. Note that the growth rate estimates $\hat{\mu}$, $\hat{r}$, $\hat{\lambda}$, and $\tilde{\alpha}$ suggest by themselves only a slow decline, if any; the population's fate is sealed by the high value of $\tilde{\sigma}^2$. We mention that the variance estimate incorporates variability from sampling as well as population fluctuations; results about hitting times reported here apply to the population *as estimated*. This point is developed further in the *Discussion* section. No qualifications to these pessimistic conclusions
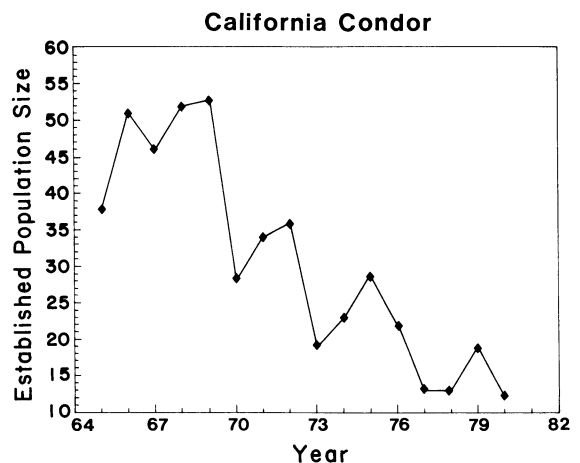
California Condor

FIG. 7. Estimated total wild population of the California Condor, 1965–1980. Data are from October surveys as listed by Wilbur (1980) and Snyder and Johnson (1985).

are found in the model diagnostic procedures; no significant outliers, influential transitions, or autocorrelation can be detected. We permit ourselves a hindsight speculation that the decision to take the field population into captivity might not have been postponed so long had such estimates been considered.

Puerto Rican Parrot

The number of breeding pairs of the highly endangered Puerto Rican Parrot has not exceeded five pairs in the last 15 yr (Wiley 1985, Snyder et al. 1987, USFWS 1987). This long-lived bird exists in the wild only in the Luquillo Forest of Puerto Rico. It breeds at 2–3 yr of age and lays 2–4 eggs. The Puerto Rican Parrot is probably the most intensively managed species reviewed in this paper. Biologists protect nest sites from predators, competitors, and human disturbance. They also rebuild nest sites, excavate new sites, remove and care for young, double clutch wild birds, and repair broken bills, wings, and feathers (USFWS 1987). As a result of this intensive management, as well as the release of captive-reared animals (by replacement in nests of juveniles that have died), the population increased from 14 in 1975 to 38 in 1989 (Fig. 8). The species would almost certainly have gone extinct without the intervention.

The data we treat here consist of the largest count of adult wild birds recorded between January and April (prebreeding period) of each year, or population estimates "by reasonable inference" made by biologists on the scene when they were convinced undercounting had occurred (Fig. 8; Snyder et al. 1987, M. Wilson, *personal communication*). In 1972, 2 birds were removed from the population of 16 birds within days of the count; thus, the 1971–1972 transition is 16–16 birds, and the 1972–1973 transition is taken as 14–16 birds.

Model parameter estimates suggest the population has favorable prospects for recovery, provided the management program continues (Tables 1 and 2). We regard this prediction as unreasonably optimistic for reasons given below. The residuals from the fitted model yield no evidence of influential transitions, autocorrelation, or outliers. We tested whether any significant change in μ occurred starting with the 1982–1983 transition, since the management efforts intensified in the early 1980s. The parameter estimates differ markedly (before: $\hat{\mu}_1 = 0.01027$; after: $\hat{\mu}_2 = 0.07808$), but the difference is not statistically significant because of the large population variability and limited number of observations ($T_{18} = -1.269$, $P = .22$). Nonetheless, a substantially higher estimate of the chance of extinction results if the transitions from 1982–1983 onward are deleted [$n_e = 1$, $n_q = 22$, $\hat{\pi} = 6.59 \times 10^{-3}$, 95% CI $= (0.00, 0.21)$].

Two ecological factors not explicitly accounted for in the model give reason for pessimism about the species' chance of survival. First, the model is merely descriptive in the sense that the cause of the popula-
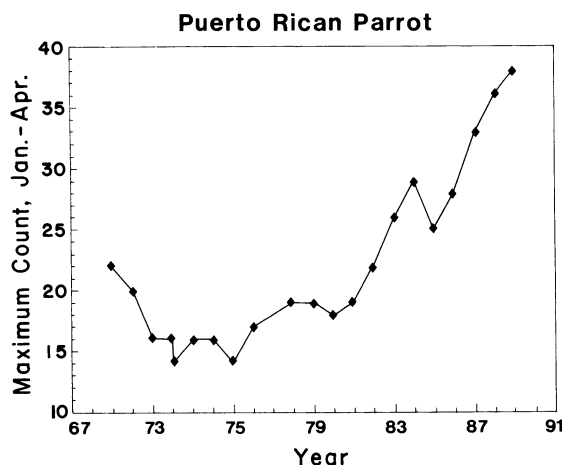

Puerto Rican Parrot

FIG. 8. Estimated total wild population of the Puerto Rican Parrot, 1969–1989. Data are from Snyder et al. (1987), supplemented by more recent counts.

tion's increase is not incorporated. This species increased because of addition of captive-reared birds, protection of juveniles, and so on; however, the breeding population has remained nearly constant. Second, environmental catastrophes are not included in the model unless such events are a regular and frequent source of variability in the time series. When this paper was being written, Hurricane Hugo had just ravaged Puerto Rico and South Carolina. A preliminary survey of the Luquillo Forest in the aftermath (M. Wilson, *personal communication*) has yielded sightings of only 8–11 birds. The birds are normally widely dispersed during the fall months, and such sightings could correspond to a population of ≈ 23 birds (M. Wilson, *personal communication*). With a starting population of 23 instead of 38, the estimated risk of declining below 10 birds increases by a factor of over 14 (Table 2). Furthermore, the values of μ and σ^2 may have changed after the storm, since extensive loss of canopy foliage cover has exposed the birds to predation from hawks. The 1989–1990 transition will no doubt be an "outlier" according to the model. The model is a supplement to, not a substitute for, basic knowledge about the population.

Palila

The Palila is a member of the Hawaiian honeycreeper family, a family known as one of the most striking examples of adaptive radiation in birds (Scott et al. 1988). This 2–2.5 cm bird weighs an average of 16 g, breeds at 1 yr of age, and lays 1–3 eggs per clutch (van Riper 1980, Berger 1981). The bird occurs only on the island of Hawaii and is restricted to portions of the upper slopes of a single mountain, Mauna Kea (Scott et al. 1984). Ongoing recovery efforts include biannual monitoring and elimination of alien species of goats and sheep, which were seriously degrading habitat on the island. Improvement of habitat quality as a result
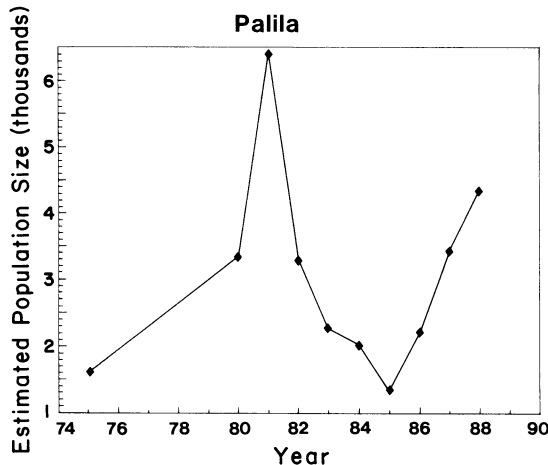
Palila

FIG. 9. Estimated total population of the Palila, 1975–1988. Data are from Scott et al. (1984), supplemented by more recent figures.

of sheep and goat removal should allow for larger population sizes more evenly distributed around the mountain. Also, plans for fire control are in place (USFWS 1978, Scott et al. 1984).

Population estimates since 1975 have varied from under 1600 to over 6000 birds (Fig. 9; Scott et al. 1984). While some of the variability is due to sampling error, fluctuations in population size account for a larger portion of the variability (Scott et al. 1984).

Model parameter estimates indicate a high propensity for population increase, but high uncertainty is present in those estimates because of the population fluctuations (Tables 1 and 2). The ML estimate of the chance of the population ever dropping below 100 birds is low, at ≈5%, but the 95% confidence interval ranges up to 50%. The chance of dropping below 1000 birds cannot be estimated with any useful degree of precision. In this situation, with a large tendency to increase but also with large stochastic fluctuations, we see the most difference between the estimates of the growth rate parameters λ and α. Recall that α discounts the small chance of the population reaching enormous sizes.

Laysan Finch

The Laysan Finch is also a Hawaiian honeycreeper. Females breed at 1 yr of age (males vary somewhat); clutch sizes range from 1 to 5 eggs and some birds produce two clutches a year (M. Morin, *personal communication*). The species is found in two populations on the Leeward islands of Hawaii in a land area of <500 ha (USFWS 1984*b*). Recovery efforts on behalf of this species include a translocation effort in 1967 by the United States Fish and Wildlife Service (Conant 1989), which resulted in the establishment of the Pearl and Hermes reef population. The originally small translocated population has increased to 500–700 birds. The establishment of the translocated population was an important step in reducing the chance of the Laysan

Finch becoming extinct through stochastic forces. Additional recovery efforts have included the elimination of feral rabbits from Laysan Island (in the Leeward islands). This alien species threatened to destroy all native vegetation. The islands were designated a national wildlife refuge early in this century, and formal recovery plans written more recently call for guarding against the accidental introduction of predators and nonnative organisms such as insects and weeds (USFWS 1984*b*). Currently, 13 Laysan Finches are in captivity.

In 22 yr of monitoring, estimates of the Leeward islands combined populations have fluctuated wildly between 5000 and 21 000 birds (Fig. 10; M. Morin, *personal communication*). As with the Palila, sampling error accounts for a part of the variability, but the fluctuations of the population itself form the largest component.

Again like the Palila, the population variability dominates the model parameter estimates (Tables 1 and 2). The extinction parameter estimates suggest that the long-term survival prospects for this species are uncertain. Though the estimated chance of dropping below 1000 birds is 1.00, great uncertainty is associated with this estimate because of the population fluctuations. More revealing is the disparity between the mean time to reach that level (over 2000 yr) and the most likely time of reaching that level (under 5 yr). Such extreme skewness of the estimated inverse Gaussian distribution indicates that explosions and catastrophes may be the rule rather than the exception in this population.

DISCUSSION

Sampling variability

When the observations in a time series are population size estimates, the data reflect sampling variability
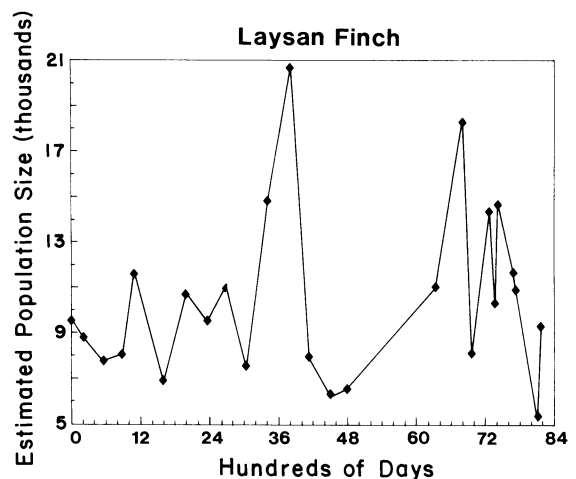
Laysan Finch

FIG. 10. Estimated total population of the Laysan Finch on Laysan Island, Hawaii, from transect samples taken at irregularly spaced time intervals, during 30 March 1969 through 16 August 1988. Data compiled by M. Morin.

as well as stochastic population fluctuations. If the sampling variation in the estimates is reasonably constant, that is, if the population sizes are estimated with relatively equal precision over time, the stochastic exponential growth model can be fitted to the series. Then the estimate of σ^2 conceptually reflects sampling variability as well as stochastic population fluctuations. The conclusions derived from fitting the model will then apply to the population (or portion of the population) as estimated, provided the model adequately describes the series. For example, in our analysis of the California Condor, $\theta(\log(12/1), \hat{\mu})$ in Table 2 is strictly described as the mean time required for the October survey population estimate to reach one bird. Such an event may or may not correspond to the population itself reaching one bird, depending on the accuracy of the population estimates.

Sampling variability can also be incorporated directly in the likelihood function for observations from a diffusion process. In general, this amounts to including an additional variance parameter. Garcia (1983), for example, included a variance component due to sampling in a diffusion model of forest growth. However, as Garcia found, one of the variance parameters becomes nearly nonidentifiable; that is, the data provide little information for its estimation. The sampling component could in principle be estimated from repeated estimates of the same population for each sampling occasion. Situations for which this is practicable are rare, though.

We mention that none of the above approaches can adequately account for sampling variation if the population sizes are estimated with widely varying precision. Thus, in analyzing the estimates of the minimal number of fully adult grizzly bear females, we are assuming that these estimates represent a reasonably constant proportion of the total population of fully adult females. This is unlikely to be the case if, for example, the aerial survey method is not standardized from year to year with respect to factors such as flying time and area covered (e.g., Seber 1982:454).

Density dependence

Perhaps the most important explicit omission in this model, and in the stochastic projection matrix formulation in general, is density dependence. Some endangered species are held in check by sheer lack of habitat. In such situations, unbridled exponential growth is hardly a likely event. The vital rates of populations near equilibrium are in balance; the fecundities and survival rates combine to produce a value of λ near 1. But to the extent that a stationary stochastic projection matrix can produce a population size appearing to fluctuate around a stable level, the diffusion model with a value of μ near 0 can describe such behavior, though not for indefinitely long periods. An example is the Laysan Finch; the population levels fluctuate but appear to have no trend up or down (Fig.

10). The Wiener-drift diffusion model adequately fits the data. The effect, however, is the introduction of large uncertainty into the parameter estimates and population predictions.

An alternate approach for modeling a population fluctuating around equilibrium is to use a diffusion process with some explicit density-dependent feedback in the infinitesimal mean. Density-dependent diffusion processes can approximate the statistical behavior of simulation models based on stochastic projection matrixes or stochastic difference equations, such as that described by Jacquez and Ginzburg (1989) and Ferson et al. (1989), or the behavior of density-dependent birth–death processes such as the model investigated by Burkey (1989). Dennis and Patil (1984) showed that an SDE version of the logistic model can serve as a general approximation for a density-dependent diffusion process having an underlying, stable equilibrium population size. Parameter estimation for this model was discussed by Dennis and Costantino (1988) and developed in more generality by Dennis (1989b); the former analyzed many examples of populations (flour beetles) fluctuating stochastically around equilibrium. Tier and Hanson (1981) derived results on the extinction (first-passage) time distribution for this logistic SDE model and others. The ramifications of using a stochastic logistic or similar model for estimating extinction-related quantities have not been investigated. Extinction parameter estimates for the Palila and the Laysan Finch might be greatly improved by such an approach.

Maximum population size

Some projection matrix models and stochastic models appearing in the conservation biology literature incorporate an upper ceiling or reflecting boundary representing a maximum population size (e.g., Suchy et al. 1985, Goodman 1987). The main effect of incorporating a ceiling is that the probability of reaching any lower boundary becomes 1; the investigations focus on properties of the extinction time (particularly the mean time to extinction). In some cases, the upper boundary is known. For instance, Suchy et al. (1985) use a value of 300 bears for the maximum grizzly population in Yellowstone, since that figure is a management target. In many cases, however, an upper boundary would have to be estimated. Such an estimate would likely take the form of an educated guess by biologists, because estimating an upper boundary from population size data in any statistically consistent fashion is extremely difficult.

We point out that extinction quantities obtained from unbounded models are often convenient approximations to quantities arising from bounded models. In the exponential growth SDE developed here, the probability of extinction (Eq. 15) is obtained as the probability of reaching x_e before reaching some upper level x_u, by taking the limit as x_u becomes large. Thus, Eq.

15 need not be regarded as implying that the population might survive eternally. Rather, Eq. 15 gives the approximate probability of reaching x_e before, say, some large upper management goal is attained. Likewise, the first-passage time, T, is approximately the waiting time required to reach x_e, given that x_e is reached before reaching the upper level.

In general, models with upper reflecting ceilings are biologically and statistically awkward. Instead, incorporation of density-dependent feedback (see above discussion, *Density dependence*) appears more promising for useful applications in conservation biology.

Types of stochastic behavior

Stochastic forces incorporated into population models have been categorized, more or less on intuitive biological grounds, as "demographic," "environmental," or "catastrophic" (May 1974, Leigh 1981, Shaffer 1981, 1987, Goodman 1987, Simberloff 1988). A fourth type of stochastic variation, "genetic" (Shaffer 1981), yields effects on population size, which would qualitatively fall into one of the above three categories (typically demographic, in the case of fluctuations in fitness from genetic drift). The fluctuations in the exponential growth SDE (Eq. 12) would appear, from our interpretation of the literature, to be of the environmental type. Unfortunately, the mathematical criteria for categorizing a given stochastic model into one of these three classes have remained vague, resulting in confusion on this point among population biologists. In particular, the type of mathematical model, such as a discrete birth–death process, continuous diffusion process, etc., does not necessarily peg the model's stochastic behavior.

We have found that the following criteria, an extension of some remarks by May (1974:33), form useful working definitions of these three types of stochastic behavior. Empirical (Pimm et al. 1988) and theoretical (Leigh 1981) studies suggest that the coefficient of variation of population size is an important component of extinction-related properties of population growth. Accordingly, let

$$\delta(t, \ n_0) = \frac{\sqrt{\text{Var}(N(t) \,|\, N(0) = n_0)}}{\text{E}[N(t) \,|\, N(0) = n_0]} \qquad (101)$$

be the coefficient of variation of the stochastic process $N(t)$. For fixed t, classify the stochastic fluctuations as demographic, environmental, or catastrophic based on whether $\delta(t, \ n_0)$ decreases to zero, becomes constant, or increase without bound, as a function of n_0.

This definition captures the spirit of the intuitive biological criteria. A model representing demographic forces essentially becomes a deterministic model for large population sizes. An example is the discrete, linear birth–death process explained in detail by Bailey (1964) and adopted by MacArthur and Wilson (1967). In addition, projection matrix simulations in which

members of each age class are subjected individually to survival–death Bernoulli trials would display demographic fluctuations. By contrast, environmental fluctuations in a model occur proportionately and are evident throughout all population sizes. By this definition, the exponential growth SDE (Eq. 12), and the logistic SDE used by Dennis and Patil (1984) and Dennis and Costantino (1988), represent environmental fluctuations. Stochastic projection matrices (see Eq. 1) would fall under this category. Note that the constant coefficient of variation could, in fact, be quite high. The variability of population size in some environmental-type models might thus appear to be rather catastrophic in magnitude; the key to the classification is that large populations fluctuate to the same proportional degree as smaller ones. Finally, models of truly catastrophic events (e.g., habitat destruction through fire) change a population by progressively larger proportions as the population size becomes larger.

A lingering problem with the proposed classification scheme is the terminology; the terms "demographic," "environmental," and "catastrophic" contain potentially misleading implications about the sources of the stochastic variability. However, the terms are widely used, and to attempt here to coin substitutes might confuse more than clarify. Another minor point is that $\delta(t, \ n_0)$, as a function of low values of n_0, might contain increasing, constant, and decreasing portions; realistic models might contain blends of stochastic forces, which vary in importance at different population sizes. The classification considers only limiting behavior as n_0 becomes large and is therefore somewhat coarse.

Time series models

An alternative approach to modeling population size is based on standard time series techniques (e.g., Pankratz 1983). The approach consists of transforming and/or differencing the data until obtaining a stationary series described by an autoregressive and/or moving average process. Such an approach can in fact yield useful information. Boyce (1987), for example, demonstrated the 10-yr cycle in the Whooping Crane data with such analyses.

The exponential growth SDE (Eq. 12) is a specific model for a nonstationary time series. Stationarity, according to the model, is achieved by log transforming and differencing the series. The resulting normal linear model for the increment data is the simplest particular case of the standard time series models. The diagnostic procedures aid in evaluating whether or not this particular model describes the data well.

The main advantages of the exponential growth SDE over other conventional time series models are as follows. First, the SDE model has a conceptual basis in the theory of age- or stage-structured population growth. Second, the time series available for real populations are often too short to take full advantage of the standard time series approach to model construction. Third, ob-

servations are frequently not collected at equal time intervals, a situation which is easily covered by the SDE model but difficult to handle by more usual time series methods. Finally, explicit and convenient results are available with the SDE model for estimating growth and extinction parameters.

That is not to say quantities such as the mean time to extinction could not be estimated from a conventional time series model. After developing such a model, one could conduct Monte Carlo simulations to estimate extinction parameters. If adequate data are available for identifying a particular time series model (or detailed age-structured model, for that matter), we recommend conducting the simulations and comparing the estimates to those described in this paper.

Management applications

The exponential growth SDE (Eq. 12) could prove to be useful in a variety of management contexts. Decisions about recovery programs for many endangered species must often be made with inadequate data available. Detailed information suitable for constructing a credible life table for a rare species seldom exists; management programs are fortunate if population size estimates, or even values of a relative population size index, are collected regularly. In such situations, the model might help biologists quantitatively assess what impact the usual, prevailing stochastic variation has on the long-term prospects for a species' survival. We caution that management policies should not be based solely on the results of the analyses described in this paper. Estimates of extinction-related quantities, as we have shown, are accompanied by large uncertainty, and effects of freak catastrophes are not accounted for in the model. However, even an order-of-magnitude assessment, or an assessment limited to the effects of environmental-type fluctuations, is better than none at all.

One crucial question regarding management of certain endangered species is whether, and if so when, to establish viable captive populations. The decision to capture some or all individuals of a species must be made before the wild population is reduced to a level having insufficient demographic and genetic diversity for a healthy founder population. The model conceivably could contribute information useful for this decision. The recovery programs for the Dusky Seaside Sparrow (*Ammodramus maritimus*), Guam Broadbill (*Myiagra freycineti*), and Kauaii Oo (*Moho braccatus*) are examples in which the debate over captive rearing raged interminably, only to have the wild population become extinct. Meaningful quantitative assessment of the species' recovery chances in these cases might have clarified earlier the urgent need for changes in management policies. There is little doubt that the 11th-hour decisions to capture the last remaining individuals of the black-footed ferret (*Mustela nigripes*) and the California Condor have provided opportunities to save these species. In the case of the California Condor, the model strongly suggests that emergency measures would have been justified some years sooner.

Another management question, in the United States at least, concerns the decision to list a species as officially "endangered" or "threatened" under the Endangered Species Act. Formal listing under the endangered or threatened categories empowers (and compels) government agencies to institute drastic measures to limit human activities that might impinge upon a species' survival chances. Predictably, in cases such as the Spotted Owl (*Strix occidentalis*) and Yellowstone grizzly bear population, biological grounds for such decisions are often discounted by political considerations (see, for instance, Simberloff 1987). Frequently lacking in the arguments over formal listing, though, are clear, defensible estimates of the risks of extinction. For example, recent upward trends in the Yellowstone grizzly population (as indexed by the estimated number of adult females; Fig. 5) have inspired optimistic proposals for removing the population from threatened status. Our estimates of the prospects for this population suggest that such proposals are premature.

CONCLUSIONS

The stochastic exponential growth model presented here combines minimum essential biological, stochastic, and statistical elements needed for assessing the chances of an endangered species' recovery. The model has a biological basis in the theory of age- or stage-structured populations. The stochastic fluctuations incorporated are of the environmental type, which are considered to pose greater problems for population persistence than those of the demographic type. The statistical interface of model with practice is well developed, so that parameter estimation, model evaluation, and forecasting are possible using only time series data on population sizes.

The model is no panacea. We do not intend for triage-style decisions about species preservation to be based solely on estimates arising from this model. The model does not incorporate some ecological factors, including freak catastrophic events and density dependence, which can be important in particular circumstances. Nonetheless, the examples we analyzed demonstrate that the model can provide useful supplemental information in a variety of situations.

Growth of endangered species is intrinsically stochastic. Accounting for stochastic forces in practice is a crucial problem in conservation biology. Such accounting will require, in addition to intensified efforts in population biology, further serious research attention to statistical questions of estimation and testing for mathematical models of population growth.

Kepler, Greg Hayward, Marie Morin, James Lewis, Philip
Dixon, and Subash Lele for their suggestions and thoughtful
comments on an earlier draft of this paper. Cameron Kepler,
Marie Morin, James Lewis, Jim Jacobi, Richard Knight, and
Marcia Wilson generously shared unpublished information
on several of the species with us; for this and other courtesies
we thank them.

LITERATURE CITED

Abramowitz, M., and I. A. Stegun. 1965. Handbook of
mathematical functions. Dover, New York, New York, USA.
Allendorf, F. W., and C. Servheen. 1986. Genetics and the
conservation of grizzly bears. Trends in Ecology and Evo-
lution 1:88–89.
Bailey, N. T. J. 1964. The elements of stochastic processes:
with applications in the natural sciences. John Wiley &
Sons, New York, New York, USA.
Basawa, I. V., and B. L. S. Prakasa-Rao. 1980. Statistical
inference for stochastic processes. Academic Press, New
York, New York, USA.
Belsley, D. A., E. Kuh, and R. E. Welsch. 1980. Regression
diagnostics: identifying influential data and sources of col-
linearity. John Wiley & Sons, New York, New York, USA.
Berger, A. J. 1981. Hawaiian birdlife. Second edition. Uni-
versity Press of Hawaii, Honolulu, Hawaii, USA.
Binkley, C. S., and R. S. Miller. 1988. Recovery of the
Whooping Crane Grus americana. Biological Conservation
45:11–20.
Boyce, M. S. 1987. Time-series analysis and forecasting of
the Aransas/Wood Buffalo Whooping Crane population.
Pages 1–9 in J. C. Lewis, editor. Proceedings of the 1985
Crane Workshop. Platte River Whooping Crane Habitat
Maintenance Trust, Grand Island, Nebraska, USA.
Boyce, M. S., and R. S. Miller. 1985. Ten-year periodicity
in Whooping Crane census. Auk 102:658–660.
Braumann, C. A. 1983a. Population growth in random en-
vironments. Bulletin of Mathematical Biology 45:635–641.
———. 1983b. Population extinction probabilities and
methods of estimation for population stochastic differential
equation models. Pages 553–559 in R. S. Bucy and J. M.
F. Moura, editors. Nonlinear stochastic problems. D. Rei-
del, Dordrecht, The Netherlands.
Burgman, M. A., H. R. Akcakaya, and S. S. Loew. 1988.
The use of extinction models for species conservation. Bi-
ological Conservation 43:9–25.
Burkey, T. V. 1989. Extinction in nature reserves: the effect
of fragmentation and the importance of migration between
reserve fragments. Oikos 55:75–81.
Capocelli, R. M., and L. M. Ricciardi. 1974. A diffusion
model for population growth in random environment. The-
oretical Population Biology 5:28–41.
Capocelli, R. M., and L. M. Ricciardi. 1979. A cybernetic
approach to population modeling. Journal of Cybernetics
9:297–312.
Chatterjee, S., and A. S. Hadi. 1988. Sensitivity analysis in
linear regression. John Wiley & Sons, New York, New York,
USA.
Cook, R. D., and S. Weisberg. 1982. Residuals and influence
in regression. Chapman and Hall, London, England.
Conant, S. 1989. Saving endangered species by transloca-

tion: are we tinkering with evolution? BioScience 38:254–
256.
Craighead, J. J., J. R. Carney, and F. C. Craighead, Jr. 1974.
A population analysis of the Yellowstone grizzly bears.
Montana Forest and Conservation Experiment Station Bul-
letin 40, University of Montana, Missoula, Montana, USA.
Dennis, B. 1989a. Allee effects: population growth, critical
density, and the chance of extinction. Natural Resource
Modeling 3:481–538.
———. 1989b. Stochastic differential equations as insect
population models. In L. McDonald, B. Manly, J. Lock-
wood, and J. Logan, editors. Estimation and analysis of
insect populations. Lecture Notes in Statistics 55:219--238.
Dennis, B., and R. F. Costantino. 1988. Analysis of steady-
state populations with the gamma abundance model: ap-
plication to Tribolium. Ecology 69:1200–1213.
Dennis, B., and G. P. Patil. 1984. The gamma distribution
and weighted multimodal gamma distributions as models
of population abundance. Mathematical Biosciences 68:187–
212.
Dennis, B., and G. P. Patil. 1988. Applications in ecology.
Chapter 12 in E. L. Crow and K. Shimizu, editors. Log-
normal distributions: theory and applications. Marcel Dek-
ker, New York, New York, USA.
Draper, N., and H. Smith. 1981. Applied regression anal-
ysis. Second edition. John Wiley & Sons, New York, New
York, USA.
Eberhardt, L. L., R. R. Knight, and B. M. Blanchard. 1986.
Monitoring grizzly bear population trends. Journal of Wild-
life Management 50:613–618.
Efron, B. 1982. The jackknife, the bootstrap, and other re-
sampling plans. Society for Industrial and Applied Math-
ematics, Philadelphia, Pennsylvania, USA.
Feldman, M. W., and J. Roughgarden. 1975. A population's
stationary distribution and chance of extinction in a sto-
chastic environment with remarks on the theory of species
packing. Theoretical Population Biology 7:197–207.
Ferson, S., L. Ginzburg, and A. Silvers. 1989. Extreme event
risk analysis for age-structured populations. Ecological
Modelling 47:175–187.
Folks, J. L., and R. S. Chhikara. 1978. The inverse Gaussian
distribution and its statistical application—a review (with
discussion). Journal of the Royal Statistical Society Series
B 40:263–289.
Garcia, O. 1983. A stochastic differential equation model
for the height growth of forest stands. Biometrics 39:1059–
1072.
Ginzburg, L. R., L. B. Slobodkin, K. Johnson, and A. G.
Bindman. 1982. Quasiextinction probabilities as a mea-
sure of impact on population growth. Risk Analysis 2:171–
181.
Goel, N. S., and N. Richter-Dyn. 1974. Stochastic models
in biology. Academic Press, New York, New York, USA.
Goodman, D. 1987. Consideration of stochastic demogra-
phy in the design and management of biological reserves.
Natural Resource Modeling 1:205–234.
Graybill, F. A. 1976. Theory and application of the linear
model. Wadsworth, Belmont, California, USA.
Heyde, C. C., and J. E. Cohen. 1985. Confidence intervals
for demographic projections based on products of random
matrices. Theoretical Population Biology 27:120–153.
Iwasa, Y., and H. Mochizuki. 1988. Probability of popu-
lation extinction accompanying a temporary decrease of
population size. Researches on Population Ecology 30:145–
164.
Jacquez, G. M., and L. Ginzburg. 1989. RAMAS: teaching
population dynamics, ecological risk assessment, and con-
servation biology. Academic Computing 4:26–27, 54–56.
Kalbfleisch, J. D. 1986. Pseudo-likelihood. Pages 324–327
in S. Kotz, N. L. Johnson, and C. B. Read, editors. Ency-

clopedia of statistical sciences. Volume 7. John Wiley & Sons, New York, New York, USA.

Karlin, S., and H. M. Taylor. 1981. A second course in stochastic processes. Academic Press, New York, New York, USA.

Keiding, N. 1975. Extinction and exponential growth in random environments. Theoretical Population Biology 8: 49–63.

Knight, R. R., and L. L. Eberhardt. 1984. Projected future abundance of the Yellowstone grizzly bear. Journal of Wildlife Management 48:1434–1438.

Knight, R. R., and L. L. Eberhardt. 1985. Population dynamics of Yellowstone grizzly bears. Ecology 66:323–334.

Koford, C. B. 1953. The California Condor. National Audubon Society Research Report 4. National Audubon Society, New York, New York, USA.

Lande, R. 1987. Extinction thresholds in demographic models of territorial populations. American Naturalist 130:624–635.

———. 1988. Genetics and demography in biological conservation. Science 241:1455–1460.

Lande, R., and S. H. Orzack. 1988. Extinction dynamics of age-structured populations in a fluctuating environment. Proceedings of the National Academy of Sciences (USA) 85:7418–7421.

Lefkovitch, L. P. 1965. The study of population growth in organisms grouped by stages. Biometrics 21:1–18.

Leigh, E. G. 1981. The average lifetime of a population in a varying environment. Journal of Theoretical Biology 90: 213–239.

Lele, S. 1991. Jackknifing linear estimating equations: asymptotic theory in stochastic processes. Journal of the Royal Statistical Society Series B 53:253–267.

Leslie, P. H. 1945. On the use of matrices in certain population mathematics. Biometrika 33:182–212.

Levinton, J. S., and L. Ginzburg. 1984. Repeatability of taxon longevity in successive foraminifera radiations and a theory of random appearance and extinction. Proceedings of the National Academy of Sciences (USA) 81:5478–5481.

Lewis, E. G. 1942. On the generation and growth of a population. Sankhya 6:93–96.

MacArthur, R. H., and E. O. Wilson. 1967. The theory of island biogeography. Princeton University Press, Princeton, New Jersey, USA.

May, R. M. 1974. Stability and complexity in model ecosystems. Second edition. Princeton University Press, Princeton, New Jersey, USA.

Mode, C. J., and M. E. Jacobson. 1987a. A study of the impact of environmental stochasticity on extinction probabilities by Monte Carlo integration. Mathematical Biosciences 83:105–125.

Mode, C. J., and M. E. Jacobson. 1987b. On estimating the critical population size for an endangered species in the presence of environmental stochasticity. Mathematical Biosciences 85:185–209.

Nedelman, J., J. A. Thompson, and R. J. Taylor. 1987. The statistical demography of Whooping Cranes. Ecology 68: 1401–1411.

Neter, J., W. Wasserman, and M. H. Kutner. 1985. Applied linear statistical models. Second edition. Richard D. Irwin, Homewood, Illinois, USA.

Nisbet, R. M., and W. S. C. Gurney. 1982. Modelling fluctuating populations. John Wiley & Sons, New York, New York, USA.

Nordheim, E. V., D. B. Hogg, and S.-Y. Chen. 1989. Leslie matrix models for insect populations with overlapping generations. In L. McDonald, B. Manly, J. Lockwood, and J. Logan, editors. Estimation and analysis of insect populations. Lecture Notes in Statistics 55:289–298.

Ogden, J. 1985. The California Condor. Pages 389–399 in

R. L. Silvestro, editor. Audubon wildlife report 1985. National Audubon Society, New York, New York, USA.

Pankratz, A. 1983. Forecasting with univariate Box-Jenkins models: concepts and cases. John Wiley & Sons, New York, New York, USA.

Pimm, S. L., H. L. Jones, and J. Diamond. 1988. On the risk of extinction. American Naturalist 132:757–785.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. 1986. Numerical recipes: the art of scientific computing. Cambridge University Press, Cambridge, England.

Ricciardi, L. M. 1977. Diffusion processes and related topics in biology. Lecture Notes in Biomathematics 14.

Rice, J. A. 1988. Mathematical statistics and data analysis. Wadsworth and Brooks/Cole, Pacific Grove, California, USA.

Richter-Dyn, N., and N. S. Goel. 1974. On the extinction of a colonizing species. Theoretical Population Biology 3: 406–433.

Rudin, W. 1964. Principles of mathematical analysis. Second edition. McGraw-Hill, New York, New York, USA.

Samson, F. B., F. Perez-Trejo, H. Salwasser, L. F. Ruggiero, and M. L. Shaffer. 1985. On determining and managing minimum population size. Wildlife Society Bulletin 13:425–433.

Schrödinger, E. 1915. Zur Theorie der Fall- und Steigversuche an Teilchen mit Brownscher Bewegung. Physikalische Zeitschrift 16:289–295.

Scott, J. M., S. Mountainspring, C. van Riper III, C. B. Kepler, J. D. Jacobi, T. A. Burr, and J. G. Giffin. 1984. Annual variation in the distribution, abundance, and habitat response of the Palila (Loxioides bailleui). Auk 101:647–664.

Scott, J. M., C. B. Kepler, C. van Riper III, and S. I. Fefer. 1988. Conservation of Hawaii's vanishing avifauna. BioScience 38:238–253.

Seber, G. A. F. 1982. The estimation of animal abundance and related parameters. Second edition. Macmillan, New York, New York, USA.

Serfling, R. J. 1980. Approximation theorems of mathematical statistics. John Wiley & Sons, New York, New York, USA.

Shaffer, M. L. 1978. Determining minimum viable population sizes: a case study of the grizzly bear (Ursus arctos L.). Dissertation. Duke University, Durham, North Carolina, USA.

———. 1981. Minimum population sizes for species conservation. BioScience 31:131–134.

———. 1987. Minimum viable populations: coping with uncertainty. Pages 69–86 in M. E. Soulé, editor. Viable populations for conservation. Cambridge University Press, Cambridge, England.

Shaffer, M. L., and F. B. Samson. 1985. Population size and extinction: a note on determining critical population sizes. American Naturalist 125:144–152.

Shimizu, K. 1988. Point estimation. Chapter 2 in E. L. Crow and K. Shimizu, editors. Lognormal distributions: theory and applications. Marcel Dekker, New York, New York, USA.

Shimizu K., and K. Iwase. 1981. Uniformly minimum variance unbiased estimation in lognormal and related distributions. Communications in Statistics, Part A—Theory and Methods 10:1127–1147.

Simberloff, D. 1987. The Spotted Owl fracas: mixing academic, applied, and political ecology. Ecology 68:766–772.

———. 1988. The contribution of population and community ecology to conservation science. Annual Review of Ecology and Systematics 19:473–511.

Slade, N. A., and H. Levenson. 1982. Estimating population growth rates from stochastic Leslie matrices. Theoretical Population Biology 22:299–308.

Snyder, N. F. R., and E. V. Johnson. 1985. Photographic censusing of the 1982–1983 California Condor population. Condor 87:1–13.

Snyder, N. F. R., J. W. Wiley, and C. B. Kepler. 1987. The parrots of Luquillo: natural history and conservation of the Puerto Rican parrot. Western Foundation of Vertebrate Zoology, Los Angeles, California, USA.

Soong, T. T. 1973. Random differential equations in science and engineering. Academic Press, New York, New York, USA.

Soulé, M. E., editor. 1986. Conservation biology: the science of scarcity and diversity. Sinauer, Sunderland, Massachusetts, USA.

———, editor. 1987. Viable populations for conservation. Cambridge University Press, Cambridge, England.

Strebel, D. E. 1985. Environmental fluctuations and extinction—single species. Theoretical Population Biology 27: 1–26.

Suchy, W. J., L. L. McDonald, M. D. Strickland, and S. H. Anderson. 1985. New estimates of minimum viable population sizes for grizzly bears of the Yellowstone ecosystem. Wildlife Society Bulletin 13:223–228.

Tier, C., and F. B. Hanson. 1981. Persistence in density dependent stochastic populations. Mathematical Biosciences 53:89–117.

Tuljapurkar, S. D. 1982a. Population dynamics in variable environments. II. Correlated environments, sensitivity analysis and dynamics. Theoretical Population Biology 21: 114–140.

———. 1982b. Population dynamics in variable environments. III. Evolutionary dynamics of r-selection. Theoretical Population Biology 21:141–165.

———. 1989. An uncertain life: demography in random environments. Theoretical Population Biology 35:227–294.

Tuljapurkar, S. D., and S. H. Orzack. 1980. Population dynamics in variable environments. I. Long-run growth rates and extinction. Theoretical Population Biology 18: 314–342.

Turelli, M. 1977. Random environments and stochastic calculus. Theoretical Population Biology 12:140–178.

Tweedie, M. C. K. 1957a. Statistical properties of the inverse Gaussian distribution I. Annals of Mathematical Statistics 28:362–377.

———. 1957b. Statistical properties of the inverse Gaussian distribution II. Annals of Mathematical Statistics 28:696–705.

USFWS. 1978. Palila recovery plan. United States Fish and Wildlife Service, Portland, Oregon, USA.

———. 1982. Grizzly bear recovery plan. United States Fish and Wildlife Service, Denver, Colorado, USA.

———. 1984a. California Condor recovery plan. United States Fish and Wildlife Service, Portland, Oregon, USA.

———. 1984b. Northwestern Hawaiian Islands passerines recovery plan. United States Fish and Wildlife Service, Portland, Oregon, USA.

———. 1985. Kirtland's Warbler recovery plan. United States Fish and Wildlife Service, Minneapolis and St. Paul, Minnesota, USA.

———. 1986. Whooping Crane recovery plan. United States Fish and Wildlife Service, Rockville, Maryland, USA.

———. 1987. Recovery plan for the Puerto Rican Parrot (Amazona vittata). United States Fish and Wildlife Service, Atlanta, Georgia, USA.

van Groenendael, J., H. de Kroon, and H. Caswell. 1988. Projection matrices in population biology. Trends in Ecology and Evolution 3:264–269.

van Riper, C., III. 1980. Observations on the breeding of the Palila (Psittirostra bailleui) of Hawaii. Ibis 122:462–475.

Walkinshaw, L. H. 1983. Kirtland's Warbler: the natural history of an endangered species. Bulletin 58. Cranbrook Institute of Science, Bloomfield Hills, Michigan, USA.

Whitmore, G. A. 1978. Discussion of the paper by Professor Folks and Dr. Chhikara. Journal of the Royal Statistical Society Series B 40:285–286.

Whitmore, G. A., and V. Sheshadri. 1987. A heuristic derivation of the inverse Gaussian distribution. American Statistician 41:280–281.

Wiemeyer, S. N., J. M. Scott, M. P. Anderson, P. H. Bloom, and C. J. Stafford. 1988. Environmental contaminants in California Condors. Journal of Wildlife Management 52: 238–247.

Wilbur, S. R. 1980. Estimating the size and trend of the California Condor population, 1965–1978. California Fish and Game 66:40–48.

Wilcox, B. A. 1986. Extinction models and conservation. Trends in Ecology and Evolution 1:46–48.

Wiley, J. W. 1985. The Puerto Rican Parrot and competition for its nest sites. Pages 213–223 in P. J. Moors, editor. Conservation of island birds. International Council for Bird Preservation, Cambridge, England.

Wright, S. J., and S. P. Hubble. 1983. Stochastic extinction and reserve size: a focal species approach. Oikos 41:466–476.

APPENDIX

Consider the inverse Gaussian cumulative distribution function (cdf) (Eq. 16). Numerical overflow can arise in evaluating (Eq. 16) at given values $\hat{\mu}$, $\hat{\sigma}^2$, x_d, and t due to the exponential function in the second term of the cdf. Note that the second term is the product of a very large number [exp($\cdot$)] and a very small number [$\Phi(\cdot)$]. The overflow problem can be circumvented if the two expressions in the product can be evaluated simultaneously.

We reparameterize Eq. 16 by defining

$$a = x_d/\sqrt{\sigma^2}, \qquad (A.1)$$
$$b = |\mu|/\sqrt{\sigma^2}, \qquad (A.2)$$
$$y = (bt - a)/\sqrt{t}, \qquad (A.3)$$
$$z = (bt + a)/\sqrt{t}, \qquad (A.4)$$

so that the inverse Gaussian cdf is

$$G(t; x_d, \mu, \sigma^2) = \Phi(y) + \exp(2ab)\Phi(-z), \qquad (A.5)$$

where $\Phi(\cdot)$ is the standard normal cdf (Eq. 7). Let $\phi(y)$ denote the derivative of $\Phi(y)$:

$$\phi(y) = (2\pi)^{-1/2}\exp(-y^2/2). \qquad (A.6)$$

Several approximations are available for evaluating $\Phi(\cdot)$ as a function of $\phi(\cdot)$ (Abramowitz and Stegun 1965). We found the following formulas to be accurate and convenient to program. If $y < 4$, then from Abramowitz and Stegun (1965: 932)

$$\Phi(y) \approx 1 - \phi(y)[d_1 q_y + d_2 q_y^2 + \ldots + d_5 q_y^5], \qquad (A.7)$$

where $q_y = 1/(1 + d_0 y)$ and $d_0 = 0.2316419$, $d_1 = 0.319381530$, $d_2 = -0.356563782$, $d_3 = 1.781477937$, $d_4 = -1.821255978$, and $d_5 = 1.330274429$. Alternatively, if $y \geq 4$ then we recommend approximating $\Phi(\cdot)$ by

$$\Phi(y) \approx 1 - \frac{\phi(y)}{y}$$
$$\cdot \left[1 - \frac{1}{y^2} + \ldots + \frac{(-1)^s 1 \cdot 3 \ldots (2s - 1)}{y^{2s}}\right], \qquad (A.8)$$

with $s = 7$ (Abramowitz and Stegun 1965:932).

Notice that

$$\phi(y) = \exp(2ab)\phi(z). \tag{A.9}$$

Expression A.9 is the key to calculating A.5, since it follows from A.9 that the product $\exp(2ab)\Phi(-z)$ can be approximated as a function of $\phi(y)$ and z. In particular, if $z < 4$ then

$$\exp(2ab)\Phi(-z) \approx \phi(y)[d_1q_z + d_2q_z^2 + \ldots + d_5q_z^5], \tag{A.10}$$

where $q_z = 1/(1 + d_0z)$, and $d_0, \ldots, d_5$ are given in Eq. A.7. For values of $z \geq 4$

$$\exp(2ab)\Phi(-z)$$
$$\approx \frac{\phi(y)}{z}\left[1 - \frac{1}{z^2} + \ldots + \frac{(-1)^s \cdot 3 \ldots (2s - 1)}{z^{2s}}\right], \tag{A.11}$$

where we suggest $s = 7$.